Office de la propriété intellectuelle du Canada

Un organisme d'Industrie Canada

Canadian Intellectual Property Office

An Agency of Industry Canada

Bureau canadien des brevets

Certification

La présente atteste que les documents ci-joints, dont la liste figure ci-dessous, sont des copies authentiques des documents déposés au Bureau des brevets.

Canadian Patent Office

Certification

This is to certify that the documents attached hereto and identified below are true copies of the documents on file in the Patent Office.

Specification and Drawings, as originally filed, with Application for Patent Serial No: **2,259,830**, on January 20, 1999, by **HSC RESEARCH AND DEVELOPMENT LIMITED.**, assignee of Daniela Rotin and Nam Pham, for "Ras Activator Nucleic Acid Molecules, Proteins and Methods of Use".

April 25, 2003

Canada

OPIC CIPO

Abstract

The invention is a RasGRF4 nucleic acid molecule and its corresponding protein which has an important role in cell signaling. This protein is regulated by Nedd4, which regulates protein stability, PDZ, which localizes the protein, and cNMP which activates it. The invention also includes biologically functional equivalent nucleic acid molecules and proteins. The invention also relates to methods of using these nucleic acid sequences and proteins in medical treatments and drug screening.

Ras Activator Nucleic Acid Molecules, Proteins and Methods of Use

FIELD OF THE INVENTION

The invention relates to isolated nucleic acid molecules encoding new Ras activator proteins identical or similar to RasGRF4. The invention also includes methods of use of RasGRF4 and the similar nucleic acid molecules and proteins for treatment of cancer and neuronal diseases, disorders and abnormal physical states.

BACKGROUND OF THE INVENTION

Activation of the Ras signaling pathway controls numerous cellular functions, most notably those regulating cell proliferation, differentiation and transformation. To date, 3 classes of Guanine Nucleotide Exchange/Releasing Factors (GEFs/GRFs) which activate Ras have been identified: (i) SOS, which binds Grb2 and connects growth factor receptors to Ras, (ii) Ras GRF1/2, which contains an IQ motif and is activated in response Ca2+/calmodulin, and (iii) RasGRP, which contains a diacylglycerol binding domain and an EF hand, and is activated by diacylglycerol and Ca2+.

Ras is involved in many aspects of cellular metabolism, so modulation of Ras activity and concentration provides a mechanism to control many cellular disease, disorders and abnormal physical states, such as cancer. None of the known classes of Ras activators have been satisfactorily modulated to control human cellular pathology. There is a clear need to identify new ways to control Ras concentration and activity.

SUMMARY OF THE INVENTION

Using an expression library screen of mouse embryonic library with the second WW domain of Nedd4 as a bait, we identified Clone 7.7, encoding about 150 amino acids, which bear 75% identity and 95% similarity to KIAA0313, a human clone (encoding an approximately 1500 amino acid protein) deposited in Genbank as part of the human genome project. The segment we isolated contained 2 PY motifs (xPPxY) which were responsible for the binding to the Nedd4-WW domain. We identified the following domains (by sequence alignment) in clone KIAA0313, and hence renamed it RasGRF4, because it represents the fourth class of Ras activators: a CDC25 homology domain (most similar to yeast CDC25 and SDC25, Ras

GRF1/2 and SOS), a PDZ domain, a cNMP binding domain (preferably cAMP-BD or cGMP-BD), a REM (Ras exchange motif) domain, a RA (Ras associating) domain, 2 PY motifs and a C terminal SAV sequence conforming to PDZ binding motif (SxV*, where * denotes STOP codon). The CDC25 of RasGRF4 domain has an approximately 70 amino acid insert, which includes a PKA phosphorylation site.

RasGRF4 schematic domain organization:

--cNMP-BD---REM---PDZ---RA---CDC25---PY-PY---SxV

We have so far demonstrated:

(i) RasGRF4 mRNA is expressed mainly in the brain (most brain regions), but also in lung.

(ii) RasGRF4 forms a stable complex with Ras *in vitro* (mainly to the nucleotide-free form of Ras and to RasGTP, but not rasGDP).

(iii) Immunoprecipitated RasGRF4 activates Ras *in vitro*. (Activation with GST-RasGRF4-CDC25 domain was variable).

(iv) Treatment of HEK-293T cells transfected with RasGRF4 with membrane permeant analogues of cAMP (8-bromo-cAMP) and cGMP (8-bromo-cGMP) leads to activation of Ras and of MAPK in RasGRF4-expressing cells but not in untransfected cells, demonstrating that these cNMP analogues can activate Ras and its downstream signaling pathway via RasGRF4. Moreover, a mutant RasGRF4 in which the cNMP-binding domain (cNMP-BD) is deleted activates Ras and MAPK constitutively, suggesting that the normal function of the cNMP-BD is to suppress the activity of the CDC25 domain, an inhibition relieved by cNMP binding or by deletion of the cNMP-BD.

(v) The PDZ domain of RasGRF4 can bind its own SAV sequence, suggesting that the protein may dimerize or fold over itself.

(vi) The protein is localized to the plasma membrane (where Ras is located), but is mislocalized in PDZ-deleted RasGRF4, suggesting that the PDZ domain is responsible for targeting/localization of RasGRF4 at the plasma membrane.

(vii) RasGRF4 transfected into mammalian cells, preferably rat2 cells, causes cellular transformation, similar to oncogenic RasV12 control.

(viii) RasGRF4 co-immunoprecipitates with Nedd4, showing that it is a target for Nedd4 ubiquitination.

Due the presence of both cNMP-BD and a PDZ domain in RasGRF4, RasGRF4 may connect G protein coupled receptors to Ras and thus to downstream signaling effectors of Ras, such as Raf-MAPK pathway, PI-3 kinase, ralGEF and possibly other effectors. G protein coupled receptors, a number of which contain a C terminal PDZ binding motif, activate adenylate cyclase via heterotrimeric G proteins, leading to increased cAMP. Thus, RasGRF4 could bind via its PDZ to these receptors at the plasma membrane and the released cAMP can directly activate (or inhibit) RasGRF4 activity and thus Ras activation. Alternatively, if cGMP is the one binding and activating (or inhibiting) RasGRF4, RasGRF may directly connect upstream activators of cGMP release (e.g. nitric oxide) to Ras. Nedd4 regulates the stability of this protein by ubiquitination, and thus suppress RasGRF4 activity by regulating its stability degradation.

The invention includes an isolated nucleic acid molecule encoding a polypeptide having RasGRF4 activity, preferably including all or part of the nucleic acid molecule of [SEQ ID NO:1]. In another embodiment, the invention includes an isolated nucleic molecule having at least 40% sequence identity to all or part of the nucleic acid molecule of [SEQ ID NO:1], wherein the nucleic acid molecule encodes a polypeptide having RasGRF4 activity.

Another embodiment is a nucleic acid molecule encoding all or part of the amino acid sequence of [SEQ ID NO:2]. The invention also includes a nucleic acid molecule that encodes all or part of a RasGRF4 polypeptide or a polypeptide having RasGRF4 activity, wherein the sequence hybridizes to the nucleic acid molecule of all or part of [SEQ ID NO:1] under high stringency conditions.

The invention includes an isolated polypeptide having RasGRF4 activity and a CDC25 domain, preferably, comprising all or part of the sequence of [SEQ ID NO:2]. The polypeptide preferably comprising at least 40% sequence identity to all or part of the polypeptide of [SEQ ID NO:2], wherein the polypeptide has RasGRF4 activity.

The invention includes a mimetic of the isolated polypeptide of any of claims 8 to 10, wherein the mimetic has RasGRF4 activity. Another aspect relates to a recombinant nucleic acid molecule comprising a nucleic acid molecule of the invention and a promoter region, operatively linked so that the promoter enhances transcription of the nucleic acid molecule in a host cell. The invention also includes a system for the expression of RasGRF4, comprising an

expression vector and a nucleic acid molecule of the invention molecule inserted in the expression vector. The invention also includes a cell transformed by the expression vector of the invention. Another aspect of the invention relates to a method for expressing polypeptide by transforming an expression host with an expression vector including and culturing the expression host.

The invention also includes a pharmaceutical composition, including all or part of the polypeptide or mimetic of the invention, and a pharmaceutically acceptable carrier, auxiliary or excipient. Another aspect of the invention relates to a RasGRF4 specific antibody targeted to a region selected from the group consisting of the C-terminus, the CDC25 domain and the PDZ domain.

The invention includes a method of medical treatment of a disease, disorder or abnormal physical state, characterized by excessive RasGRF4 expression, concentration or activity, comprising administering a product that reduces or inhibits RasGRF4 polypeptide expression, concentration or activity. The invention also includes a method of medical treatment of a disease, disorder or abnormal physical state, characterized by inadequate RasGRF4 expression, concentration or activity, comprising administering a product that increases RasGRF4 polypeptide expression, concentration or activity.

BRIEF DESCRIPTION OF THE DRAWINGS

Preferred embodiments of the invention will be described in relation to the drawings in which:

Figure 1. Domain organization of Rat Nedd4.

Figure 2. Protein sequence of Clone 7.7, the homolog of human clone KIAA0313.

Figure 3A. Schematic Diagram of RasGRF4.

Figure 3B. cDNA and amino acid sequence of RasGRF4 (KIAA0313).

Figure 4A. Protein sequence alignment of CDC25 domains from several RasGEF/GRF including RasGRF4.

Figure 4B. Comparison of CDC25 domain of RasGRF4 with RasGRF2 revealing the insert in RasGRF4.

Figure 5. Protein sequence of alignment of Ras GRF4-REM domain.

Figure 6A. Overall structure comparison between RasGRF4 and other known mammalianGRFs/GEFs which activate Ras.

Figure 6B. An example of the most well known Ras signaling pathway.

Figure 7. Protein sequence alignment of RasGRF4-PDZ domain.

Figure 8. Protein sequence alignment of RasGRF4-cCAMP binding domain.

Figure 9. Protein sequence alignment of RAsGRF4-RA domain.

Figure 10. Tissue Distribution of RasGRF4.

Figure 11. Co-precipitation of endogenous Nedd4 in Hek 293T cells by a GST-fusion protein of the C-terminal last 150 aa of RasGRF4 which contains the two PY motifs .

Figure 12. Co-immunoprecipitation of RasGRF4 with endogenous Nedd4 in Hek 293T cells transiently transfected with Flag-tagged RasGRF4.

Figure 13. Method used for the *in vitro* GEF assay.

Figure 14. *In vitro* GEF assay using immunoprecipitated full-length RasGRF4 demonstrating activation of Ras by RasGRF4.

Figure 15. RasGRF4 forms stable complex with GST-Ras in vitro.

Figure 16. GRF4 induces foci formation in Rat2 fibroblasts.

Figure 17. GST-fusion protein of RasGRF4-PDZ domain binds full-length RasGRF4 expressed in Hek 293T cells.

Figure 18. Biotinylated peptide of the last 15 aa sequence of RasGRF4 containing a putative PDZ-binding motif (SAV*) binds full-length RasGRF4.

Figure 19. (a) Human clone KIAA0313 full DNA sequence [SEQ ID NO:1] and amino acid sequence [SEQ ID NO:2]; (b) Clone 7.7 DNA sequence [SEQ ID NO:3] and amino acid sequences [SEQ ID NOS:4,5,6].

Figure 20. Plasma membrane localization of RasGRF4.

DETAILED DESCRIPTION OF THE INVENTION

Identification and characterization of RasGRF4

We compared the sequence of mouse Clone 7.7, which we isolated, to Genbank sequences and identified K1AA0313 as the human homologue. KIAA0313 was a protein of unknown function in Genbank database when we first reviewed its sequence. We characterized this protein experimentally as a Ras-specific guanine-nucleotide releasing factor (Ras-GRF) and renamed this protein as RasGRF4.

The invention includes RasGRF4 nucleic acid molecules and molecules having sequence identity or which hybridize to the RasGRF4 sequence which encode a protein capable of activating Ras (preferred percentages for sequence identity are described below) as well as vectors including these molecules. The invention also includes RasGRF4 or proteins having sequence identity (preferred percentages described below) or which are capable of activating Ras. The nucleic acid molecules and proteins of the invention may be from lung, brain or the neuronal system and they may be isolated from a native source, synthetic or recombinant. The invention includes RasGRF4 or proteins having sequence identity (preferred percentages described below) which are capable of activating Ras, as prepared by the processes described in this application.

This GRF represents a fourth class of RasGRFs. Fig. 3 is a schematic diagram of RasGRF4. Taken together, the structural features of RasGRF4 show a multifunctional role that involves regulation of several aspects of cell physiology, including cell proliferation, morphology, membrane transport, cell survival and cellular transformation. Our finding that RasGRF4 overexpression causes cell transformation shows that RasGRF4 is oncogenic. RasGRF4 expression, concentration and activity may be manipulated in methods of medical treatment of excessive cell proliferation, such as in cancer.

The RasGRF4 DNA encodes an approximately 1500 amino acid protein containing various domains and motifs. RasGRF4 has several unique domains which are not found in other mammalian RasGRFs and which are important for unique regulation of its activity. These unique domains include a PDZ domain, a putative cNMP-binding domain (cAMP-BD or cGMP-BD) and a RA domain, two PY motifs, a coiled-coil motif and a C-terminal SXV motif, conforming to the PDZ binding motif.

RasGRF4 activity and effects on Ras

RasGRF4 is activated by distinct signaling pathways that involve a G-coupled receptor signaling pathway (Fig. 19). RasGRF4 can be activated by a G-protein coupled receptor via an association of RasGRF4-PDZ domain and its binding motif present in many such receptor. This activation process may depend on the activation state of the receptor. Binding of RasGRF4 to such a receptor leads to activation of RasGRF4 as a result of conformational changes or membrane recruitment of RasGRF4 (or both). In one of the aspects of the inventions, activation of a G-coupled receptor leads to elevation of cAMP which modulates RasGRF4 activity by directly binding to RasGRF4-cAMP-BD. The SAV* motif of RasGRF4 can be involved in an intramolecular interaction with RasGRF4-PDZ domain and this interaction may have regulatory roles in RasGRF4 activity. Likewise, this motif can bind to other PDZ-containing proteins associating with the plasma membrane. RasGRF4 binds preferentially to nucleotide-free and GTP-bound Ras. The RA domain of RasGRF4 mediates RasGRF4 binding to Ras-GTP. In so doing, RasGRF4 functions as a downstream Ras effector. Nedd4 interacts with RasGRF4 through WW domain-PY motif interaction and ubiquitinates RasGRF4 and targets it for degradation.

RasGRF4 domains and motifs

PY-motifs

RasGRF4 contains two PY-motifs near the C-terminus which bind to Nedd4-WW domains leading to its identification as a Nedd4-WW domain interacting protein in the expression library screen. Preferable protein hybridization conditions use TBS-Tween (about: 137 mM NaCl, 27 mM KCl, 25 mM tris, pH 8.0, 0.1% Tween 20). The screen used to identify Clone 7.7 was based on protein:protein interactions (i.e. a labeled GST Nedd4-WW domain protein was used as a probe to screen an expression library. cDNA of the library was induced to express proteins. Washes were done with TBS-Tween). These conditions can be used in a method to identify other GRF proteins similar to RasGRF4 which preferably have RasGRF4 activity or similar activity.

CDC25 Domains

RasGRF4 harbours a central catalytic region called CDC25 domain, named for the prototypic Ras activator in Saccharomyces cerevisiae (21), from which the putative function of RasGRF4 was deduced.

CDC25 domains have been shown to catalyze guanine-nucleotide exchange/release activity on Ras family GTPases. The CDC25 of RasGRF4 is 48-52% similar to those of yeast CDC25, SOS and RasGRF/RasGRF2. Fig. 4 shows the alignment of CDC25 domains from various proteins including RasGRF4. From the mutagenesis studies of yeast CDC25, several conserved arginine residues were proposed to be critical for its activity (22). These conserved arginine residues are also conserved in RasGRF4. Similar to CDC25, SDC25, RasGRF1/2 and SOS, RasGRF4 contains blocks of highly conserved sequences (Fig. 4A) which were recently demonstrated, based on the tertiary structure of SOS bound to Ras, to play a critical role in the activity of the CDC25 domain towards Ras (23). However, unique to RasGRF4, the RasGRF4-CDC25 domain also contains an insert (about 70 amino acids) not found in SOS, RasGRF2 or other RasGRF3 (Fig. 3B).

Ras Exchange of Motif Domain

RasGRF4 also has a REM (Ras exchange motif) domain (24) which is present in all known mammalian RasGRFs. Fig. 5 shows the alignment of REM domains from several proteins including RasGRF4. Mammalian RasGRFs all share this REM domain which is likely important for their activities. Recently, it was reported that the REM domain of SOS contributes to the activity of the CDC25 domain by stabilizing the active structure of the catalytic region (23).

Diacylglycerol Binding Domain, EF Hands, Calcium Binding Motif

As shown in Fig. 6A, each mammalian RasGRF has its own unique domains which are important for regulation of its activity. Specifically, SOS was shown to be activated by various growth factors, a process involving binding of activated receptor-tyrosine kinase to Grb2-SH2 domain and Grb2-SH3 domain to the proline-rich region of SOS - (25). RasGRF1 and RasGRF2 were shown to be activated by elevation of intracellular calcium, a process involving the binding of Ca2+-bound calmodulin to the IQ motif present in these RasGRFs (23, 26). RasGRP harbours a DAG (diacylglycerol) binding domain and a pair of EF hands, a Ca2+ binding motif and accordingly, it was shown to be activated by elevated level of DAG and calcium (27). These unique domains allow RasGRFs to activate Ras in response to distinct signaling pathways. The small GTPase Ras controls the MAPK pathway, (as well as PI-3 kinase, ralGEF and likely other effectors). In so doing, Ras exerts its effects on many cellular processes such as cellular proliferation and differentiation (Fig.6B).

PDZ Domains

PDZ (PSD95/Dlg/ZO-1) domains, also known as DHR (Disc-large homology region) or GLGF domains (conserved stretch of amino acids in the domain) are 80 -100 amino acid protein-protein interaction modules which are found in membrane-associating proteins and intracellular signaling proteins (Ref. 28). PDZ domains are important for membrane targeting, clustering of receptors/channels and forming scaffold of networks of signaling proteins at the plasma membrane. Examples include PSD-95 which binds the NMDA receptors, as well as the InaD which binds to the TRP, components of photo-transduction cascades in the Drosophila eyes (29-30). PDZ domains were shown to bind to C-terminal three or four residues in a sequence specific context. One class of PDZ domains, including those of Disc-large protein, were shown to bind to C-terminal Valine residue in a context of S/T x V* (* denotes a stop codon). While other classes of PDZ domains were shown to bind C-terminal three residues with hydrophobic or aromatic side chains (31). The alignment of PDZ domains of several proteins including RasGRF4 is given in Fig. 7. The PDZ domain of RasGRF4 is similar to a class of PDZ domains binding S/T x V* motif. RasGRF4 itself has such a motif (SAV*) at its C-terminus (Fig. 3), so there is likely to be interaction between RasGRF4-PDZ domain and its own putative PDZ-binding motif.

cNMP Binding Domain

RasGRF4 has a cNMP-binding domain that preferably binds cAMP or cGMP. It shares 50% sequence similarity to that of the regulatory subunits of PKA. Fig. 8 shows the alignment of cNMP-binding domains. Since a conformational change is often accompanied by binding of cNMP to a protein, RasGRF4 activity may be regulated by conformational changes. By having a cAMP-binding domain, RasGRF4 may be involved in a G-coupled receptor pathway and connect this pathway to the Ras signaling pathway. Many G-protein coupled receptors contain putative PDZ-binding motifs which bind and regulate activities of PDZ- domain containing proteins. Having both a PDZ domain and a putative cAMP binding domain, RasGRF4 is likely regulated by a G-coupled receptor system coupling to the adenylyl cyclase enzyme. Alternatively, if cGMP is the compound binding and activating (or inhibiting) RasGRF4, RasGRF may directly connect upstream activators of cGMP release (e.g. nitric oxide) to Ras.

Ras Associating Domain

RasGRF4 also has a RA (Ras associating) domain. This type of domain was initially identified in two Ras effector proteins, including RalGDS and AF-6/Canoe, and later in numerous putative Ras binding proteins. RA domains have been assumed to bind to Ras-GTP and the solved tertiary structure of RalGDS-RA domain was found to be similar to that of the Ras binding domain of Raf kinase which binds to Ras-GTP (32). However, recent evidence suggests that not all putative RA domains bind to Ras-GTP. The alignment of RA domains from several proteins including RasGRF4 is given in Fig. 9.

PEST Sequences, coil-coil and PY motifs

In addition to the above domains, RasGRF4 has two PEST sequences which associate with unstable proteins. RasGRF4 also has a coiled-coil region which likely participates in protein-protein interaction through interactions of multiple amphipathic alpha helices (33). The PY motifs serve as attachment sites for the Nedd4-WW domain, thereby facilitating ubiquitination and degradation of RasGRF4.

Functionally equivalent nucleic acid molecules

The invention includes nucleic acid molecules that are functional equivalents of all or part of the sequence in [SEQ ID NO:1]. (A nucleic acid molecule may also be referred to as a DNA sequence or nucleotide sequence in this application. All these terms have the same meaning as nucleic acid molecule and may be used to refer, for example, to a cDNA, complete gene or a gene fragment. The intended meaning will be clear to a person skilled in the art.) Functionally equivalent nucleic acid molecules are DNA and RNA (such as genomic DNA, cDNA, synthetic DNA, and mRNA nucleic acid molecules), that encode peptides, proteins, and polypeptides having the same or similar RasGRF4 activity as the RasGRF4 polypeptide shown in [SEQ ID NO:2]. Functionally equivalent nucleic acid molecules can encode peptides, polypeptides and proteins that contain a region having sequence identity to a region of a RasGRF4 polypeptide or more preferably to the entire RasGRF4 polypeptide. The CDC25 is a preferred region because it is the central catalytic region. The invention includes nucleic acid molecules that have a region with sequence identity to the CDC25 coding region of [SEQ ID NO:1] which is represented by about nucleotide no. 2194 (2131+63) to nucleotide no. 3082 (preferred percentages of identity are below). The invention includes nucleic acid molecules about: <1000 nucleotides (preferably about 888 nucleotides), < 1500 nucleotides, <2000 nucleotides, <3000 nucleotides or <5000 nucleotides which encode a

region having sequence identity to the CDC25 coding region and having CDC25 activity or CDC25-like activity.

Identity is calculated according to methods known in the art. The Clustal W program (preferably using default parameters) [Thompson, JD et al., Nucleic Acid Res. 22:4673-4680.], described below, is most preferred. For example, if a nucleic acid molecule (called "Sequence A") has 90% identity to a portion of the nucleic acid molecule in [SEQ ID NO:1], then Sequence A will preferably be identical to the referenced portion of the nucleic acid molecule in [SEQ ID NO:1], except that Sequence A may include up to 10 point mutations, such as substitutions with other nucleotides, per each 100 amino acids of the referenced portion of the nucleic acid molecule in [SEQ ID NO:1]. Mutations described in this application preferably do not disrupt the reading frame of the coding sequence. Nucleic acid molecules functionally equivalent to the RasGRF4 sequences can occur in a variety of forms as described below.

Nucleic acid molecules may encode conservative amino acid changes in RasGRF4 polypeptide. The invention includes functionally equivalent nucleic acid molecules that encode conservative amino acid changes within a RasGRF4 amino acid sequence and produce silent amino acid changes in RasGRF4.

Nucleic acid molecules may encode non-conservative amino acid substitutions, additions or deletions in RasGRF4 polypeptide. The invention includes functionally equivalent nucleic acid molecules that make non conservative amino acid changes within the RasGRF4 amino acid sequence in [SEQ ID NO:2]. Functionally equivalent nucleic acid molecules include DNA and RNA that encode peptides, polypeptides and proteins having non-conservative amino acid substitutions (preferably substitution of a chemically similar amino acid), additions, or deletions but which also retain the same or similar RasGRF4 activity as the RasGRF4 polypeptide shown in [SEQ ID NO:2]. The DNA or RNA can encode fragments or variants of RasGRF4. Fragments are useful as imminogens and in immunogenic compositions (U.S. Patent No. 5,837,472). The RasGRF4 or RasGRF4 -like activity of such fragments and variants is identified by assays as described below. Fragments and variants of RasGRF4 encompassed by the present invention should preferably have at least about 40%, 60%, 80% or 95% sequence identity or preferably at least about 96%, 97%, 98%, 99%, 99.5%, 99.9% or more preferably at least about 99.95% sequence identity to the naturally occurring RasGRF4 nucleic acid molecule (preferably measured between the coding region of the KIAA0313 sequence nucleotides 63 to 4562), or a region of the sequence, such as the coding sequence or one of the conserved domains of the nucleic acid molecule, without being

identical to the sequence in [SEQ ID NO:1]. These sequences preferably encode all the RasGRF4 domains and motifs described above. One or more domain or motif may be omitted to obtain desired activity. The CDC25 domain is preferably conserved in the nucleic acid molecule and polypeptide in order to preserve RasGRF4 activity. Sequence identity is preferably measured with the Clustal W program (preferably using default parameters) [Thompson, JD et al., Nucleic Acid Res. 22:4673-4680.]. In another embodiment, the Gap program may be used. The algorithm of Needleman and Wunsch (1970 J. Mol. Biol. 48:443-453) is used in the Gap program. BestFit may also be used to measure sequence identity. It aligns the best segment of similarity between two sequences. Alignments are made using the local homology algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2:482-489. Most preferably, 1, 2, 3, 4, 5, 5-10, 10-15, 15-25, 25-50, 50-100 or 100-600 nucleotides are modified. One would be able to make more changes to the nucleotide and amino acid sequences (such as substitutions, deletions) in regions outside of the conserved regions of RasGRF4 described above.

Nucleic acid molecules functionally equivalent to the RasGRF4 in [SEQ ID NO:1] will be apparent from the following description. For example, the sequence shown in [SEQ ID NO:1] may have its length altered by natural or artificial mutations such as partial nucleotide insertion or deletion, so that when the entire length of the coding sequence within [SEQ ID NO:1], is taken as 100%, the functional equivalent nucleic acid molecule preferably has a length of about 60-120% thereof, more preferably about 80-110% thereof. Fragments may be less than 60%.

Nucleic acid molecules containing partial (usually 80% or less, preferably 60% or less, more preferably 40% or less of the entire length) natural or artificial mutations so that some codons in these sequences code for different amino acids, but wherein the resulting polypeptide retains the same or similar RasGRF4 activity as that of a naturally occurring RasGRF4 polypeptide. The mutated DNAs created in this manner should preferably encode a polypeptide having at least about 40%, preferably at least about 60%, at least about 80%, and more preferably at least about 90% or 95%, and most preferably at least about 97%, 98%, 99%, 99.5%, 99.9%, or 99.95% sequence identity to the amino acid sequence of the RasGRF4 polypeptide in [SEQ ID NO:2]. Sequence identity is preferably assessed by the Clustal W program.

Since the genetic code is degenerate, the nucleic acid sequence in [SEQ ID NO:1] is not the only sequence which may code for a polypeptide having RasGRF4 activity. This

invention includes nucleic acid molecules that have the same essential genetic information as the nucleic acid molecule described in [SEQ ID NO:1] or a domain or motif of this region. Nucleic acid molecules (including RNA) having one or more nucleic acid changes compared to the sequences described in this application and which result in production of a polypeptide shown in [SEQ ID NO:2] are within the scope of the invention.

Other functional equivalent forms of RasGRF4 -encoding nucleic acids can be isolated using conventional DNA-DNA or DNA-RNA hybridization techniques. Thus, the present invention also includes nucleic acid molecules that hybridize to one or more of the sequences in [SEQ ID NO:1] or its complementary sequence, and that encode expression for peptides, polypeptides and proteins exhibiting the same or similar activity as that of the RasGRF4 polypeptide produced by the DNA in [SEQ ID NO:1] or its variants. Such nucleic acid molecules preferably hybridize to the sequence in [SEQ ID NO:1] under moderate to high stringency conditions (see Sambrook et al. Molecular Cloning: A Laboratory Manual, Most Recent Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). High stringency washes have low salt (preferably about 0.2% SSC), and low stringency washes have high salt (preferably about 2% SSC). A temperature of about 37 °C or about 42 °C is considered low stringency, and a temperature of about 50-65 °C is high stringency. The invention also includes a method of identifying nucleic acid molecules encoding a RasGRF4 activator polypeptide (preferably a mammalian polypeptide), including contacting a sample containing nucleic acid molecules including all or part of [SEQ ID NO:1] (preferably at least about 15 or 30 nucleotides of [SEQ ID NO:1]) under moderate or high stringency hybridization conditions and identifying nucleic acid molecules which hybridize to the nucleic acid molecules including all or part of [SEQ ID NO:1].) Similar methods are described in U.S. Patent No. 5,851,788 which is Incorporated by reference in its entirety.

The invention also includes methods of using all or part of the nucleic acid molecules which hybridize to all or part of [SEQ ID NO:1], for example as probes or in assays to identify antagonists or inhibitors of the polypeptides produced by the nucleic acid molecules (described below). The invention also includes methods of using nucleic acid molecules having sequence identity to the RasGRF4 nucleic acid molecule (as described below) in similar methods.

The invention also includes a nucleic acid molecule detection kit including, preferably in a suitable container means or attached to a surface, a nucleic acid molecule of the invention encoding RasGRF4 or a polypeptide having RasGRF4 activity and a detection reagent (such as a detectable label). Other variants of kits will be apparent from this description and

teachings in patents such as U.S. Patent Nos. 5,837,472 and 5,801,233 which are Incorporated by reference in their entirety.

For example, Hybridization solution 1 is low stringency: about: >50 % formamide, >5X denhardt's, >1% SDS, >5X SSC, >42 °C; Hybridization solution 2 is high stringency: about: >1% BSA, >1mM EDTA, >0.5 M NaHPO4, pH 7.2, >7% SDS, >65 °C. A preferable high stringency wash consists of about: >0.2 X SSC, >0.1% SDS. A preferable low stringency wash has about: >2XSSC, >0.1% SDS).

The present invention also includes nucleic acid molecules that hybridize to genomic DNA, cDNA, or synthetic DNA molecules that encode the amino acid sequence of the RasGRF4 polypeptide, or genetically degenerate forms, under salt and temperature conditions equivalent to those described in this application, and that encode a peptide, polypeptide or polypeptide that has the same or similar activity as the RasGRF4 polypeptide. In a preferred embodiment, the invention includes DNA that hybridizes to all or part of the CDC25 coding region of [SEQ ID NO:1] which is represented by about nucleotide no. 2194 (2131+63) to nucleotide no. 3082, under moderate to high stringency conditions.

A nucleic acid molecule described above is considered to have a function substantially equivalent to the RasGRF4 nucleic acid molecules of the present invention if the polypeptide produced by the nucleic acid molecule has RasGRF4 activity. A polypeptide has RasGRF4 activity if it can activate Ras. Activation of Ras is shown where a polypeptide is active in catalyzing guanine-nucleotide exchange on small GTPase Ras using the *in vitro* GEF assay.

Production of RasGRF4 in eukaryotic and prokaryotic cells

The nucleic acid molecules of the invention may be obtained from a cDNA library. The nucleotide molecules can also be obtained from other sources known in the art such as expressed sequence tag analysis or *in vitro* synthesis. The DNA described in this application (including variants that are functional equivalents) can be introduced into and expressed in a variety of eukaryotic and prokaryotic host cells. A recombinant nucleic acid molecule for the RasGRF4 contains suitable operatively linked transcriptional or translational regulatory elements. Suitable regulatory elements are derived from a variety of sources, and they may be readily selected by one with ordinary skill in the art (Sambrook, J, Fritsch, E.E. & Maniatis, T. (Most Recent Edition). Molecular Cloning: A laboratory manual. Cold Spring Harbor Laboratory Press. New York; Ausubel et al. (Most Recent Edition) Current Protocols in Molecular Biology, John Wiley & Sons, Inc.). For example, if one were to upregulate the expression of the nucleic acid molecule, one could insert a sense sequence and the

appropriate promoter into the vector. Promoters can be inducible or constitutive, environmentally - or developmentally-regulated, or cell - or tissue-specific. Transcription is enhanced with promoters known in the art for expression. The CMV and SV40 promoters are commonly used to express desired polypeptide in mammalian cells. Other promoters known in the art may also be used (many suitable promoters and vectors are described in the applications and patents referenced in this application).

If one were to downregulate the expression of the nucleic acid molecule, one could insert the antisense sequence and the appropriate promoter into the vehicle. The nucleic acid molecule may be either isolated from a native source (in sense or antisense orientations), synthesized, or it may be a mutated native or synthetic sequence or a combination of these.

Examples of regulatory elements include a transcriptional promoter and enhancer or RNA polymerase binding sequence, a ribosomal binding sequence, including a translation initiation signal. Additionally, depending on the vector employed, other genetic elements, such as selectable markers, may be incorporated into the recombinant molecule. Other regulatory regions that may be used include an enhancer domain and a termination region. The regulatory elements may be from animal, plant, yeast, bacterial, fungal, viral, avian, insect or other sources, including synthetically produced elements and mutated elements.

In addition to using the expression vectors described above, the polypeptide may be expressed by inserting a recombinant nucleic acid molecule in a known expression system derived from bacteria, viruses, yeast, mammals, insects, fungi or birds. The recombinant molecule may be introduced into the cells by techniques such as *Agrobacterium tumefaciens*-mediated transformation, particle-bombardment-mediated transformation, direct uptake, microinjection, coprecipitation, transfection and electroporation depending on the cell type. Retroviral vectors, adenoviral vectors, Adeno Associated Virus (AAV) vectors, DNA virus vectors and liposomes may be used. Suitable constructs are inserted in an expression vector, which may also include markers for selection of transformed cells. The construct may be inserted at a site created by restriction enzymes.

In one embodiment of the invention, a cell is transfected with a nucleic acid molecule of the invention inserted in an expression vector to produce cells expressing a polypeptide encoded by the nucleic acid molecule.

Another embodiment of the invention relates to a method of transfecting a cell with a nucleic acid molecule of the invention, inserted in an expression vector to produce a cell

expressing the RasGRF4 polypeptide or other polypeptide of the invention. The invention also relates to a method of expressing the polypeptides of the invention in a cell. A preferred process would include culturing a cell including a recombinant DNA vector including a nucleic acid molecule encoding RasGRF4 (or another nucleic acid molecule of the invention) in a culture medium so that the polypeptide is expressed. The process preferably further includes recovering the polypeptide from the cells or culture medium.

Probes

The invention also includes oligonucleotide probes made from the cloned RasGRF4 nucleic acid molecules described in this application or other nucleic acid molecules of the invention, such as Clone 7.7 (see materials and methods section). The probes may be 15 to 30 nucleotides in length and are preferably at least 30 or more nucleotides. A preferred probe is at least 15 nucleotides of RasGRF4 in [SEQ ID NO:1] or the Clone 7.7 sequence. The invention also includes at least 30 consecutive nucleotides of [SEQ ID NO:1] or the Clone 7.7 sequence. The probes are useful to identify nucleic acids encoding RasGRF4 peptides, polypeptides and polypeptides other than those described in the application, as well as peptides, polypeptides and polypeptides functionally equivalent to RasGRF4. The oligonucleotide probes are capable of hybridizing to the sequence shown in [SEQ ID NO:1] under stringent hybridization conditions. A nucleic acid molecule encoding a polypeptide of the invention may be isolated from other organisms by screening a library under moderate to high stringency hybridisation conditions with a labeled probe. The activity of the polypeptide encoded by the nucleic acid molecule is assessed by cloning and expression of the DNA. After the expression product is isolated the polypeptide is assayed for RasGRF4 activity as described in this application.

Functionally equivalent RasGRF4 nucleic acid molecules from other cells, or equivalent RasGRF4 -encoding cDNAs or synthetic DNAs, can also be isolated by amplification using Polymerase Chain Reaction (PCR) methods. Oligonucleotide primers, such as degenerate primers, based on [SEQ ID NO:2] can be prepared and used with PCR and reverse transcriptase (E. S. Kawasaki (1990), In Innis et al., Eds., PCR Protocols, Academic Press, San Diego, Chapter 3, p. 21) to amplify functional equivalent DNAs from genomic or cDNA libraries of other organisms. The oligonucleotides can also be used as probes to screen cDNA libraries.

Functionally equivalent peptides, polypeptides and proteins

The present invention includes not only the polypeptides encoded by the sequences of the invention, but also functionally equivalent peptides, polypeptides and proteins that exhibit the same or similar RasGRF4 polypeptide activity. A polypeptide is considered to possess a function substantially equivalent to that of the RasGRF4 polypeptide if it has RasGRF4 activity. Functionally equivalent peptides, polypeptides and proteins include peptides, polypeptides and proteins that have the same or similar protein activity as RasGRF4 when assayed, i.e. they are able to activate Ras. A polypeptide has RasGRF4 activity if it is active in catalyzing guanine-nucleotide exchange on small GTPase Ras using the in-vitro GEF assay. (Where only one or two of the terms peptides, polypeptides and proteins is referred to, it will be clear to one skilled in the art whether the other types of amino acid sequences also would be useful.)

These peptides, polypeptides and proteins can contain a region or moiety exhibiting sequence identity to a corresponding region or moiety of the RasGRF4 polypeptide described in the application, but this is not required as long as they exhibit the same or similar RasGRF4 activity.

Identity refers to the similarity of two polypeptides or proteins that are aligned so that the highest order match is obtained. Identity is calculated according to methods known in the art, such as the Clustal W program. For example, if a polypeptide (called "Sequence A") has 90% identity to a portion of the polypeptide in [SEQ ID NO:2], then Sequence A will be identical to the referenced portion of the polypeptide in [SEQ ID NO:2], except that Sequence A may include up to 10 point mutations, such as substitutions with other amino acids, per each 100 amino acids of the referenced portion of the polypeptide in sequence (a) in [SEQ ID NO:2]. Peptides, polypeptides and proteins functional equivalent to the RasGRF4 polypeptides can occur in a variety of forms as described below.

Peptides, polypeptides and proteins biologically functional equivalent to RasGRF4 polypeptide include amino acid sequences containing amino acid changes in the RasGRF4 sequence. The functional equivalent peptides, polypeptides and proteins have at least about 40% sequence identity, preferably at least about 60%, at least about 75%, at least about 80%, at least about 90% or at least about 95% sequence identity, to the naturally RasGRF4 polypeptide or a corresponding region. More preferably, the functional equivalent peptides, polypeptides and proteins have at least about 97%, 98%, 99%, 99.5%, 99.9% or 99.95% sequence identity to the naturally occurring RasGRF4 polypeptide or a region of the sequence

(such as one of the conserved domains of the polypeptide), without being identical to the sequence in [SEQ ID NO: 2] . "Sequence identity" is preferably determined by the Clustal W program. Most preferably, 1, 2, 3, 4, 5, 5-10, 10-15, 15-25 or 25-50 amino acids are modified. The sequences preferably include all the RasGRF4 domains and motifs described above. One or more domain or motif may be omitted to obtain desired activity. The CDC25 domain is preferably conserved in the polypeptide in order to preserve RasGRF4 activity. Structurally conserved regions 1, 2 and 3 (Fig. 4A) are critical for CDC25 structure and activity. Preferably, conserved amino acids in these regions would not be altered. One would be able to make more changes to the amino acid sequences in regions outside of the conserved regions of RasGRF4. The CDC25 region of the polypeptide includes amino acid no. 712 to amino acid no. 1006 (preferred percentages of identity are below). The invention includes polypeptides about: <350 amino acids (preferably about 294 amino acids), < 500 amino acids, < 750 amino acids, < 1000 amino acids, <1250 amino acids, <1500 amino acids or < 2000 amino acids which have sequence identity to the CDC25 region and have CDC25 activity or CDC25-like activity (preferably Ras activation).

The invention includes peptides, proteins or proteins which retain the same or similar activity as all or part of RasGRF4. Such peptides preferably consist of at least 5 amino acids. In preferred embodiments, they may consist of 6 to 10, 11 to 15, 16 to 25 or 26 to 50, 50 to 150, 150 to 250, 250 to 500, 500 to 750 or 750 to 1250 amino acids of RasGRF4. Fragments of the RasGRF4 polypeptide can be created by deleting one or more amino acids from the N-terminus, C-terminus or an internal region of the polypeptide (or combinations of these), so long as the fragments retain the same or similar RasGRF4 activity as all or part of the RasGRF4 polypeptide disclosed in the application. These fragments can be generated by restriction nuclease treatment of an encoding nucleic acid molecule. Alternatively, the fragments may be natural mutants of the RasGRF4. Fragments of the polypeptide may be used in an assay to identify compounds that bind the polypeptide. Methods known in the art may be used to identify agonists and antagonists of the fragments.

Variants of the RasGRF4 polypeptide may also be created by splicing. A combination of techniques known in the art may be used to substitute, delete or add amino acids. For example, a hydrophobic residue such as methionine can be substituted for another hydrophobic residue such as alanine. An alanine residue may be substituted with a more hydrophobic residue such as leucine, valine or isoleucine. An aromatic residue such as phenylalanine may be substituted for tyrosine. An acidic, negatively charged amino acid such

as aspartic acid may be substituted for glutamic acid. A positively charged amino acid such as lysine may be substituted for another positively charged amino acid such as arginine. Modifications of the polypeptides of the invention may also be made by treating a polypeptide of the invention with an agent that chemically alters a side group, for example, by converting a hydrogen group to another group such as a hydroxy or amino group.

Peptides having one or more D-amino acids are contemplated within the invention. Also contemplated are peptides where one or more amino acids are acetylated at the N-terminus. Those skilled in the art recognize that a variety of techniques are available for constructing peptide mimetics (i.e. a modified peptide or polypeptide or protein) with the same or similar desired biological activity as the corresponding polypeptide of the invention but with more favorable activity than the polypeptide with respect to characteristics such as solubility, stability, and/or susceptibility to hydrolysis and proteolysis. See for example, Morgan and Gainor, *Ann. Rep. Med. Chem.*, 24:243-252 (1989).

The invention also includes hybrid nucleic acid molecules and peptides, for example where a nucleic acid molecule from the nucleic acid molecule of the invention is combined with another nucleic acid molecule to produce a nucleic acid molecule which expresses a fusion peptide. A preferred fusion polypeptide includes all or part of the active CDC25 Domain of RasGRF4. One or more of the other domains of RasGRF4 described in this application could also be used to make fusion polypeptides. For example, a nucleotide domain from a molecule of interest may be ligated to all or part of a nucleic acid molecule encoding RasGRF4 polypeptide (or a molecule having sequence identity) described in this application. Fusion nucleic acid molecules and peptides can also be chemically synthesized or produced using other known techniques. The invention includes a nucleic acid molecule encoding a fusion polypeptide or a recombinant vector including the sequence of [SEQ ID NO:1] or [SEQ ID NO:3]. The invention also includes a fusion polypeptide including the sequence of [SEQ ID NO:2] or a polypeptide encoded by [SEQ ID NO:3].

The variants preferably retain the same or similar RasGRF4 activity as the naturally occurring RasGRF4. The RasGRF4 activity of such variants can be assayed by techniques described in this application and known in the art.

Variants produced by combinations of the techniques described above but which retain the same or similar RasGRF4 activity as naturally occurring RasGRF4 are also included in the invention (for example, combinations of amino acid additions, deletions, and substitutions).

Fragments and variants of RasGRF4 encompassed by the present invention preferably have at least about 40% sequence identity, preferably at least about 60%, 75%, 80%, 90% or 95% sequence identity, to the naturally occurring polypeptide, or corresponding region or moiety. Most preferably, the fragments have at least about 97%, 98% or 99%, 99.5%, 99.9% or 99.99% sequence identity to the naturally occurring RasGRF4 polypeptide, or corresponding region. Sequence identity is preferably measured with the Clustal W.

The invention also includes fragments of the polypeptides of the invention which do not retain the same or similar activity as the complete polypeptides but which can be used as a research tool to characterize the polypeptides of the invention.

Enhancement of RasGRF4 polypeptide activity

The activity of the RasGRF4 polypeptide is increased by carrying out selective site-directed mutagenesis. Using protein modeling and other prediction methods, we characterize the binding domain and other critical amino acid residues in the polypeptide that are candidates for mutation, insertion and/or deletion. A DNA plasmid or expression vector containing the RasGRF4 nucleic acid molecule or a nucleic acid molecule having sequence identity is preferably used for these studies using the U.S.E. (Unique site elimination) mutagenesis kit from Pharmacia Biotech or other mutagenesis kits that are commercially available, or using PCR. Once the mutation is created and confirmed by DNA sequence analysis, the mutant polypeptide is expressed using an expression system and its activity is monitored. This approach is useful not only to enhance activity, but also to engineer some functional domains for other properties useful in the purification or application of the polypeptides or the addition of other biological functions. It is also possible to synthesize a DNA fragment based on the sequence of the polypeptides that encodes smaller polypeptides that retain activity and are easier to express. It is also possible to modify the expression of the cDNA so that it is induced under desired environmental conditions or in response to different chemical inducers or hormones. It is also possible to modify the DNA sequence so that the polypeptide is targeted to a different location. All these modifications of the DNA sequences presented in this application and the polypeptides produced by the modified sequences are encompassed by the present invention.

Pharmaceutical compositions

The RasGRF4 nucleic acid molecule or its polypeptide and functional equivalent nucleic acid molecules or polypeptides are also useful when combined with a carrier in a

pharmaceutical composition. Suitable examples of vectors for RasGRF4 are described above. The compositions are useful when administered in methods of medical treatment of a disease, disorder or abnormal physical state characterized by insufficient RasGRF4 expression or inadequate levels or activity of RasGRF4 polypeptide by increasing expression, concentration or activity. The invention also includes methods of medical treatment of a disease, disorder or abnormal physical state characterized by excessive RasGRF4 expression or levels or activity of RasGRF4 polypeptide, for example by administering a pharmaceutical composition including a carrier and a vector that expresses RasGRF4 antisense DNA. Cancer is one example of a disease which can be treated by antagonizing RasGRF4. An agent that upregulates RasGRF4 gene expression or RasGRF4 polypeptide activity may be combined with a carrier to form a pharmaceutical composition. An agent that downregulates RasGRF4 expression or RasGRF4 polypeptide activity may be combined with a carrier to form a pharmaceutical composition.

The pharmaceutical compositions of this invention are used to treat patients having degenerative diseases, disorders or abnormal physical states such as cancer and diseases associated with nervous system function. For example, cancer can be treated by antagonizing RasGRF4, by blocking CDC25 activity. The following U.S. patents deal with the use of compounds that modulate Ras in order to treat diseases, disorders or abnormal physical states: 5856439, 5852034, 5843941, 5840683, 5807853, 5801175, 5789438, 5776902, 5756528, 5712280, 5710171, 5672611, 5668171, 5663193, 5661128, 5627202, 5624936, 5585359, 5582995, 5576293, 5571835, 5567729, 5536750, 5523456, 5491164, 5480893, 5468733, 5238922, 5185248, 5523456, 5491164, 5480893, 5468733, 5238922 and 5185248 which are incorporated by reference in their entirety. The following WIPO PCT patent applications disclose the use of compounds that modulate Ras in order to treat diseases: WO9857990, WO9805786, WO9828980, WO9815556, WO9857970, WO9857964, WO9857963, WO9857949, WO9857948, WO9857947, WO9857946, WO9849194, WO9811106, WO9811098, WO9811097, WO9809641, WO9804545, WO9721820, WO9857950 and WO9737678 which are incorporated by reference in their entirety. Many of these patents and applications describe inhibition of Ras to treat excessive cell proliferation and cancer. The patents and applications disclose research techniques to identify compounds which inhibit Ras or compounds that regulate Ras.

The pharmaceutical compositions can be administered to humans or animals by methods such as tablets, aerosol administration, intratracheal instillation and intravenous injection in

methods of medical treatment involving upregulating or downregulating RasGRF4 gene or polypeptide to upregulate or downregulate Ras activity. Dosages to be administered depend on patient needs, on the desired effect and on the chosen route of administration.

Nucleic acid molecules and polypeptides may be introduced into cells using *in vivo* delivery vehicles such as liposomes. They may also be introduced into these cells using physical techniques such as microinjection and electroporation or chemical methods such as coprecipitation or using liposomes.

The pharmaceutical compositions can be prepared by known methods for the preparation of pharmaceutically acceptable compositions which can be administered to patients, and such that an effective quantity of the nucleic acid molecule or polypeptide is combined in a mixture with a pharmaceutically acceptable vehicle. Suitable vehicles are described, for example in Remington's Pharmaceutical Sciences (Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pa., USA).

On this basis, the pharmaceutical compositions could include an active compound or substance, such as a RasGRF4 nucleic acid molecule or polypeptide, in association with one or more pharmaceutically acceptable vehicles or diluents, and contained in buffered solutions with a suitable pH and isoosmotic with the physiological fluids. The methods of combining the active molecules with the vehicles or combining them with diluents is well known to those skilled in the art. The composition could include a targeting agent for the transport of the active compound to specified sites within tissue.

Administration of RasGRF4 nucleic acid molecule

Since persons suffering from disease, disorder or abnormal physical state can be treated by either up or down regulation of RasGRF4, gene therapy to increase or reduce RasGRF4 expression is useful to modify the development/progression of disease. For example, to treat cancer, RasGRF4 could be modulated to suppress Ras activity (inhibiting RasGRF4 prevents Ras activation).

The invention also includes methods and compositions for providing gene therapy for treatment of diseases, disorders or abnormal physical states characterized by insufficient RasGRF4 expression or inadequate levels or activity of RasGRF4 polypeptide (see the discussion of phamaceutical compositions, above) involving administration of a pharmaceutical composition of the invention. The invention also includes methods and compositions for providing gene therapy for treatment of diseases, disorders or abnormal

physical states characterized by excessive RasGRF4 expression or levels of activity of RasGRF4 polypeptide involving administration of a pharmaceutical composition.

The invention includes methods and compositions for providing a nucleic acid molecule encoding RasGRF4 or functional equivalent nucleic acid molecule to the cells of an individual such that expression of RasGRF4 in the cells provides the biological activity or phenotype of RasGRF4 polypeptide to those cells (preferably Ras activation). Sufficient amounts of the nucleic acid molecule are administered and expressed at sufficient levels to provide the biological activity or phenotype of RasGRF4 polypeptide to the cells. For example, the method can preferably involve a method of delivering a nucleic acid molecule encoding RasGRF4 to the cells of an individual having a disease, disorder or abnormal physical state, comprising administering to the individual a vector comprising DNA encoding RasGRF4. The method may also relate to a method for providing an individual having a disease, disorder or abnormal physical state with biologically active RasGRF4 polypeptide by administering DNA encoding RasGRF4. The method may be performed *ex vivo* or *in vivo*. Methods and compositions for administering RasGRF4 (including in gene therapy) are explained, for example, in U.S. Patent Nos. 5,672,344, 5,645,829, 5,741,486, 5,656,465, 5,547,932, 5,529,774, 5,436,146, 5,399,346 and 5,670,488, 5,240,846 which are incorporated by reference in their entirety.

The method also relates to a method for producing a stock of recombinant virus by producing virus suitable for gene therapy comprising DNA encoding RasGRF4. This method preferably involves transfecting cells permissive for virus replication (the virus containing the nucleic acid molecule) and collecting the virus produced.

The invention also includes methods and compositions for providing a nucleic acid molecule encoding an antisense sequence to RasGRF4 or a Nedd4 nucleic acid molecule sequence to the cells of an individual such that expression of the sequence prevents RasGRF4 biological activity or phenotype or reduces RasGRF4. The methods and compositions can be used *in vivo* or *in vitro*. Sufficient amounts of the nucleic acid molecule are administered and expressed at sufficient levels to reduce the biological activity or phenotype of RasGRF4 polypeptide in the cells. Similar methods as described in the preceding paragraph may be used with appropriate modifications.

The methods and compositions can be used *in vivo* or *in vitro*. The invention also includes compositions (preferably pharmaceutical compositions for gene therapy). The compositions include a vector containing RasGRF4. Nedd4 or a functional equivalent

molecule or antisense DNA. The carrier may be a pharmaceutical carrier or a host cell transformant including the vector. Vectors known in the art include adenovirus, adeno associated virus (AAV), herpesvirus vectors, such as vaccinia virus vectors, and plasmids. The invention also includes packaging cell lines that produce the vector. Methods of producing the vector and methods of gene therapy using the vector are also included with the invention.

The invention also includes a transformed cell, such as a brain cell or a lung cell containing the vector and recombinant RasGRF4 nucleic acid molecule antisense sequence, Nedd4 or a functionally equivalent molecule.

Heterologous expression of RasGRF4

Expression vectors are useful to provide high levels of polypeptide expression. Cell cultures transformed with the nucleic acid molecules of the invention are useful as research tools particularly for studies of RasGRF4 interactions with Ras. Novel pathways to activate Ras are identified. Cell cultures are used in overexpression and research according to numerous techniques known in the art. For example, a cell line (either an immortalized cell culture or a primary cell culture) may be transfected with a vector containing a RasGRF4 nucleic acid molecule (or molecule having sequence identity) to measure levels of expression of the nucleic acid molecule and the activity of the nucleic acid molecule and polypeptide. A polypeptide of the invention may be used in an assay to identify compounds that bind the polypeptide. Methods known in the art may be used to identify agonists and antagonists of the polypeptides. One may obtain cells that do not express RasGRF4 endogenously and use them in experiments to assess ectopoic RasGRF4 nucleic acid molecule expression. Experimental groups of cells may be transfected with vectors containing different types of RasGRF4 nucleic acid molecules (or nucleic acid molecules having sequence identity to RasGRF4 or fragments of RasGRF4 nucleic acid molecule) to assess the levels of polypeptide produced, its functionality and the phenotype of the cells produced. Other expression systems can also be utilized to overexpress the RasGRF4 in recombinant systems. The polypeptides are also useful for *in vitro* analysis of RasGRF4 activity. For example, the polypeptide produced can be used for microscopy or X-ray crystallography studies, and the tertiary structure of individual domains may be analyzed by NMR spectroscopy.

Experiments may be performed with cell cultures or in vivo to identify polypeptides that bind to different domains of RasGRF4. One could also target cNMP to block upstream activators or inhibitors. Nedd4 binding to RasGRF4 can be studied. For example, Nedd4

binding could be blocked to study the effects on RasGRF4 stability. Another example is blocking the PDZ domain to prevent membrane localization of RasGRF4. Similar approaches could be taken to study other polypeptide domains or motifs.

Preparation of antibodies

The RasGRF4 polypeptide is also useful as an antigen for the preparation of antibodies that can be used to purify or detect other RasGRF4-like polypeptides. To recognize the polypeptide: preferably target to the C-terminus. To block activity: preferably target to the CDC25 domain. To block membrane targeting: preferably target to the PDZ domain.

We have already generated polyclonal antibodies against the C-terminal 150 amino acids of RasGRF4 which is a unique region. Monoclonal and polyclonal antibodies are prepared according to the description in this application and techniques known in the art. For examples of methods of the preparation and uses of monoclonal antibodies, see U.S. Patent Nos. 5,688,681, 5,688,657, 5,683,693, 5,667,781, 5,665,356, 5,591,628, 5,510,241, 5,503,987, 5,501,988, 5,500,345 and 5,496,705 which are incorporated by reference in their entirety. Examples of the preparation and uses of polyclonal antibodies are disclosed in U.S. Patent Nos. 5,512,282, 4,828,985, 5,225,331 and 5,124,147 which are incorporated by reference in their entirety. Antibodies recognizing RasGRF4 can be employed to screen organisms or tissues containing RasGRF4 polypeptide or RasGRF4-like polypeptides. The antibodies are also valuable for immuno-purification of RasGRF4 or RasGRF4-like polypeptides from crude extracts.

An antibody (preferably the antibody described above) may be used to detect RasGRF4 or a similar polypeptide, for example, by contacting a biological sample with the antibody under conditions allowing the formation of an immunological complex between the antibody and a polypeptide recognized by the antibody and detecting the presence or absence of the immunological complex whereby the presence of RasGRF4 or a similar polypeptide is detected in the sample. The invention also includes compositions preferably including the antibody, a medium suitable for the formation of an immunological complex between the antibody and a polypeptide recognized by the antibody and a reagent capable of detecting the immunolgical complex to ascertain the presence of RasGRF4 or a similar polypeptide. The invention also includes a kit for the *in vitro* detection of the presence or absence of RasGRF4 or a similar polypeptide in a biological sample, wherein the kit preferably includes an antibody, a medium suitable for the formation of an immunological complex between the antibody and a

polypeptide recognized by the antibody and a reagent capable of detecting the immunological complex to ascertain the presence of RasGRF4 or a similar polypeptide in a biological sample. Further background on the use of antibodies is provided, for example in U.S. Patent Nos. 5,695,931 and 5,837,472 which are Incorporated by reference in their entirety.

Diagnostic test

In many cancers, Ras is aberrantly expressed or is mutated. It is likely that in some cancers, RasGRF4 is mutated as well, so RasGRF4 is useful as a screening tool for the detection of cancer or to monitor its progression. For example, RasGRF4 may be sequenced to determine if a cancer-causing mutation is present. Levels of RasGRF4 may also be measured to determine whether RasGRF4 is upregulated.

Screening for agonists and antagonists of RasGRF4 and enhancers and inhibitors of RasGRF4 polypeptide

As described above, RasGRF4 is useful in a pharmaceutical preparation to treat cancer and other diseases disorders and abnormal physical states. Nedd4 (preferably all or part of Nedd4, such as the RasGRF4 binding domain of Nedd4) is one agent which reduces RasGRF4 activity. cAMP and cGMP are agents which increase RasGRF4 activity. RasGRF4 is also useful as a target. Modulation of RasGRF4 expression is commercially useful for identification and development of drugs to inhibit and/or enhance RasGRF4 function directly. Such drugs would preferably be targeted to any of the following sites: CDC25 domain, PDZ domain, cNMP-BD. Chemical libraries are used to identify pharmacophores which can specifically interact with RasGRF4 either in an inhibitory or stimulatory mode. The RasGRF4 targets that would be used in drug design include the CDC25 domain, in order to inhibit its catalytic activity. For example, nucleotide analogues which stabilize the Ras-analogue complex, thus preventing replacement of the nucleotide analogue by Ras, could interfere with activation of RasGRF4. Similarly, other compounds directed against the binding site of Ras on RasGRF4 could be useful as well. The insert in the CDC25 domain in RasGRF4 is unique and is useful as a target. The PDZ domain is necessary for proper localization of RasGRF4 to the plasma membrane and is useful as a target. The cNMP binding domain is useful to disconnect RasGRF4 from upstream signaling. The invention also includes methods of screening a test compound to determine whether it antagonizes or agonizes RasGRF4 polypeptide activity. The invention also includes methods of screening a test compound to determine whether it induces or inhibits RasGRF4 nucleic acid molecule expression.

In a preferred embodiment, the invention includes an assay for evaluating whether test compounds are capable of acting as agonists or antagonists for RasGRF4, or a polypeptide having RasGRF4 functional activity, including culturing cells containing DNA which expresses RasGRF4, or a polypeptide having RasGRF4 activity so that the culturing is carried out in the presence of at least one compound whose ability to modulate RasGRF4 activity (preferably Ras activating activity or CDC25 domain activity) is sought to be determined and thereafter monitoring the cells for either an increase or decrease in the level of RasGRF4 or RasGRF4 activity. Other assays (as well as variations of the above assay) will be apparent from the description of this invention and techniques such as those disclosed in U.S. Patent No. 5,851,788, which is incorporated by reference in its entirety. For example, the test compound levels may be either fixed or increase.

Localization of RasGRF4

i) Tissue distribution of RasGRF4

To show tissue distribution of RasGRF4, mouse RasGRF4 specific probes were used to probe a Rat multiple tissue mRNA blot (Clonetech). Two messages, of 8.5 and 7.5 Kb, are present in rat brain; the 8.5 Kb message is also present in rat lung (Fig.10). We determine the polypeptide's distribution in neuronal tissue. The finding of RasGRF4 message in rat brain is consistent with the fact that its cDNA was initially isolated from a human brain cDNA library. Using human RasGRF4 specific probes on the human brain multiple region mRNA blots (Clonetech), RasGRF4 messages (8.5 and 7.5 Kb) are found widespread (Fig. 10). The two messages may correspond to splicing variants or isoforms of RasGRF4. In comparison, SOS is ubiquitously expressed, whereas RasGRF1, RasGRF2 and RasGRP are expressed primarily in the brain (23,26,27). We detect RasGRF4 polypeptides in cell lines using known techniques.

ii) Characterization of Nedd4-RasGRF4 interaction

Since mouse RasGRF4 was isolated from the expression library screen using Nedd4-WW2 domain as a probe, further characterization of their interaction was studied.

A GST-fusion protein of polypeptide corresponding to the last 150 amino acid of RasGRF4 (about the same length as the partial amino acid sequence isolated from the screen), containing the two PY motifs of RasGRF4, was generated and used in a pull-down experiment. Nedd4 is endogenously expressed in Hek 293T cells and can be detected in 293T lysates using Nedd4 antibodies (Fig. 11). When 293T lysates were incubated with

agarose beads bound to GST or GST-fusion protein of the PY-containing polypeptide, Nedd4 was found to bind specifically to this polypeptide, showing that the two PY motifs of RasGRF4 are sufficient to interact with full-length Nedd4.

The interaction between Nedd4 and RasGRF4 was also demonstrated in living cells by co-immunoprecipitation. Flag-epitope tagged RasGRF4 was constructed in a mammalian expression vector (pCMV5). The co-immunoprecipitation experiment was performed using endogenous Nedd4 and transiently transfected Flag-tagged RasGRF4 in Hek 293T cells. First, Flag-tagged RasGRF4 was immunoprecipitated from transfected lysates using anti-Flag gel affinity (Sigma). When this immunocomplex containing RasGRF4 was resolved on SDS-PAGE and subsequently immunoblotted with Nedd4 antibodies, Nedd4 was detected in this immunocomplex. However, Nedd4 was not found in the immunocomplex that did not have RasGRF4 when lysates of cells transfected with empty vector were used (Fig. 12). Therefore, Nedd4 is co-immunoprecipitating with RasGRF4, showing that they interact in living cells.

RasGRF4 also contains PEST sequences. RasGRF4 is an unstable protein which is ubiquitinated by Nedd4 and targeted for degradation via the ubiquitin-dependent proteolytic pathway. We perform a ubiquitination assay to show that RasGRF4 is ubiquitinated protein using the protocol described in Ref. 34.

iii) In-vitro guanine nucleotide exchange activities

RasGRF4 has a RasGRF(GEF) activity / function. To show its GEF activity, we performed in-vitro GEF assays. The schematic outline of the in-vitro GEF assay protocol (described in Ref. 24) is given in Fig. 13. Briefly, GST-Ras was added alone (tubes 1 and 2) or along with GST-CDC25, or immunoprecipitated full-length of RasGRF4 (tubes 3 and 4). All tubes contained assay mixture including cold GTP and P32 alpha GTP. The exchange reactions were stopped at the indicated times. The stopped reaction mixtures were passed through nitrocellulose filters which were then washed with stop buffer to separate bound and unbound nucleotides. Filters were dried and then quantified by scintillation counting to determine bound CPM. The labeled nucleotides trapped on the washed filters were assumed to be Ras-associated. The difference in bound CPM over 30 minute period was determined for reactions where GST-Ras was added alone (it is the difference in bound CPM between tubes 1 and 2) and where GST-Ras was added with a GEF (it is the difference in bound CPM between tubes 3 and 4). The former is the basal level of GTP-binding to Ras and the later is usually increased several folds over the basal activity if the indicated GEF is active.

Using the GEF assay described above, the immunoprecipitated full-length RasGRF4 was shown to be active on Ras (Fig. 14). Similar levels of GEF activity were also observed for the immunoprecipitated full-length RasGRF2 used as a positive control in this assay.

We perform in-vitro GEF assays using GST-CDC25 of RasGRF4 to show that this domain is sufficient for activity.

iv) In-vitro interaction of RasGRF4 with Ras

In order to show that RasGRF4 can form a stable complex with Ras in vitro, and which nucleotide-bound forms of Ras it binds preferentially, an in-vitro pull-down experiment was performed as follows: Lysates of 293T cells transiently transfected with Flag-tagged RasGRF4 were incubated with agarose beads bound to either GST alone or GST-Ras of different nucleotide-bound states. Beads were washed and resolved on SDS-PAGE and subsequently immunoblotted with anti-Flag antibodies to detect Flag-tagged RasGRF4. The results showed that RasGRF4 bound specifically to Ras as it failed to bind to GST alone. However, it bound to Ras differentially, depending on the nucleotide-bound states of Ras. RasGRF4 bound strongly to EDTA-treated Ras (EDTA chelates Mg2+ which is important for binding of nucleotides to Ras, thus keeps Ras in nucleotide-free form) and Ras-GTP, but bound weakly to Ras-GDP (Fig. 15). In similar experiments, RasGRF2 was shown to bind only to EDTA-treated Ras (23).

v) Activation of Ras and MAPK by cAMP and cAMP analogues:

Treatment of HEK-293T cells transfected with RasGRF4 with membrane permeant analogues of cAMP (8-bromo-cAMP) and cGMP (8-bromo-cGMP) leads to activation of Ras and of MAPK in RasGRF4-expressing cells but not in untransfected cells, demonstrating that these cNMP analogues activate Ras and its downstream signaling pathway via RasGRF4. Moreover, a mutant RasGRF4 in which the cNMP-binding domain (cNMP-BD) is deleted activates Ras and MAPK constitutively, showing that the normal function of the cNMP-BD is to suppress the activity of the CDC25 domain, an inhibition relieved by cNMP binding or by deletion of the cNMP-BD.

vi) Transformation assay

The small GTPase Ras functions as a molecular switch in cells by switching between its inactive form when it is bound to GDP and its active form when it is bound to GTP. RasGRFs activate Ras by promoting nucleotide exchange from GDP (inactive) to GTP (active) on Ras. Active Ras activates the MAPK pathway and other signaling pathways to control

normal cellular events such as cellular proliferation and differentiation. However, when Ras activity can not be deactivated as in the case of mutant oncogenic Ras, Ras becomes oncogenic and its transforming ability is the underlying mechanism of cellular transformation and is the cause of many human cancers (Ref 41-44). Several signaling proteins upstream and downstream of Ras, either controlling the activity of Ras or carrying out Ras effects, were also shown to be oncogenic.

We showed that RasGRF4 can transform cells overexpressing this protein. Transformation assays were performed using Rat 2 fibroblasts, a suitable cell type for this assay. Rat 2 cells were transiently transfected with empty vector, RasGRF4 construct, or mutant RasV12 construct (a transforming form of Ras used as a positive control). After transfection, cells were cultured over a period of three weeks with routine changes of media, and were routinely examined for morphology changes under a light microscope. Fig. 16 shows the result of the assay. Rat 2 cells transfected with empty vector grew at moderate rate and maintained a monolayer state of normal saturation density, as seen with non-transfected cells. In contrast, Rat 2 cells transfected with the RasGRF4 construct grew faster, achieved much higher saturation density as compared to cells transfected with empty vector; more importantly, RasGRF4 induced foci formation in these transfected cells. A focus is the site where a single transformed cell proliferates and forms a prolific mass of transformed cells; foci formation shows a loss of cell-cell contact inhibition, a hallmark of cellular transformation. A similar phenotype was also observed with Rat 2 cells transfected with RasV12 construct. The finding that RasGRF4 induces foci formation in Rat 2 fibroblasts shows that RasGRF4 is oncogenic as well as highlights the physiological importance of this protein.

vii) PDZ domain of RasGRF4 interacts with its own PDZ-binding motif, SAV*

RasGRF4 harbours a PDZ domain and a putative PDZ-binding motif in context of SAV* and thus, it is involved in potential intramolecular interaction or intermolecular homotypic interaction.

The following experiment indicates that the PDZ domain of RasGRF4 binds to its own SAV* motif and thus gives rise to either intramolecular interaction or intermolecular homotypic interaction. A GST-fusion protein of RasGRF4-PDZ domain (GST-PDZ) was generated and used in a pull-down experiment. Lysates of 293T cells transfected with Flag-tagged full-length RasGRF4 were incubated with agarose beads bound to GST alone or GST-PDZ. Beads were washed and resolved on SDS-PAGE and subsequently immunoblotted with anti-Flag antibodies to detect bound RasGRF4. As shown in Fig. 17, the full-length RasGRF4 binds

specifically to GST-PDZ, showing that the interaction is mediated by binding of GST-PDZ to the SAV* motif present in the full-length RasGRF4. Furthermore, in a similar pull-down experiment, the strep-tavidin agarose beads bound to biotinylated peptide corresponding the last 15 amino acids of RasGRF4 (therefore, containing the SAV* motif) were shown to bind to the full-length RasGRF4 also, thus suggesting again an interaction between the PDZ domain and the SAV* motif of RasGRF4 (Fig. 18).

viii) Immunofluorescence studies / Localization

We determined that RasGRF4 exhibits plasma-membrane staining and is localized at the plasma membrane where Ras, its substrate, is located. This plasma membrane localization is mediated by the PDZ domain because the protein is localized diffusely in the cytosol upon deletion of the PDZ domain.

Additional Examples

a) Activation of Ras by RasGRF4:

We have already demonstrated that full-length RasGRF4 is active in catalyzing guanine-nucleotide exchange on small GTPase Ras using in-vitro GEF assay. As mentioned earlier, RasGRF4 has a REM domain which is present in all mammalian RasGRFs and therefore, we believe that RasGRF4 is a Ras-specific GRF. We test RasGRF4 activity on other small GTPases of Ras family (Ral, Rap and so on) and those of Rho family (Rho, Rac and Cdc42) and show that RasGRF4 is a Ras specific GRF.

We also determine whether the RasGRF4-CDC25 domain is necessary and sufficient for its activity. First, we construct a mutant RasGRF4 construct lacking the CDC25 domain which can be expressed in mammalian cells and used in in-vitro GEF assays. This mutant construct, along with the full-length RasGRF4 which was already shown to be active on Ras, is measured for its activity or loss of activity. Furthermore, a GST-fusion protein of RasGRF4-CDC25 domain is generated and used in an in-vitro GEF assay to show that RasGRF4-CDC25 domain is sufficient for the RasGRF4 activity. RasGRF4 lacking the CDC25 domain will lose its ability to modulate Ras.

Concurrently, we measure the GEF activity on Ras of RasGRF4 on Ras in living cells, using the method described in Ref 35. This method employs a GST-fusion protein of Ras-binding domain (RBD) of Raf kinase (Raf is an immediate downstream kinase of Ras in MAPK pathway). Raf-RBD binds to Ras-GTP (active Ras) and thus is useful to assay levels of active Ras in cells. GST-RBD is incubated with lysates of cells transfected with RasGRF4 or empty

vector. Active Ras in lystates is precipitated by GST-RBD beads and detected by anti-Ras antibodies on Western blot. In cells transfected with RasGRF4, we show more active Ras being pulled down by GST-RBD. This in vivo Ras activation assay also allows us to test effects of various treatments to cells of RasGRF4 activity.

We characterize the activation mechanisms of RasGRF4 and the signaling pathways employed by RasGRF4 from these in vivo Ras activation assays. For instance, since RasGRF4 has a cNMP-binding domain (cAMP-BD or cGMP-BD) we showed that cAMP or cAMP analogues activate RasGRF4. We construct a GST-fusion protein of this cAMP-BD in order to demonstrate its in-vitro binding affinity towards cAMP or cAMP using protocol previously described in Ref. 36.

We determine the roles of individual domains of RasGRF4 in Ras activation. We construct various mutant RasGRF4 constructs lacking individual domains which are tested for their activities on Ras using both in-vitro GEF assay and in vivo Ras activation assay.

The small GTPase Ras controls the MAPK pathway and exerts its effects on cellular processes primarily through this pathway. MAPK is a downstream kinase of Ras and thus, Ras activation leads to MAPK activation (Fig.6A). Therefore, we show the RasGRF4 effects on MAPK activation using assays in which levels of active MAPK in cells is determined using antibodies recognizing phosphorylated (active) MAPK.

b) Transforming ability of RasGRF4:

We already showed that RasGRF4 induces Rat 2 fibroblasts to form foci which are indicative of a loss of cell-cell contact inhibition. We use a mutant RasGRF4 construct lacking the catalytic domain which is therefore enzymatically inactive in the transformation assays alongside with the full-length RasGRF4 construct, in order to show that the CDC25 domain is necessary for the observed phenotype.

A loss of cell-cell contact inhibition and anchorage-independent growth are the two hallmarks of cellular transformation. These two properties underline the mechanism of tumor formation and metastasis. The oncogenic Ras and other oncogenes were already shown to exhibit these two transforming properties. We perform soft-agar assays to measure RasGRF4 anchorage-independent growth in Rat 2 cells transfected with RasGRF4.

We study the transforming ability of RasGRF4 in living animals. Tumor-formation assay is performed in nude mice ectopically injected with RasGRF4-induced transformed Rat2 cells.

c) The activation mechanisms and signaling pathways employed by RasGRF4:

Although all known mammalian RasGRFs are activated by different signals arising from distinct signaling pathways (Fig.6B), they all appear to employ similar activation mechanisms once they are recruited to the plasma membrane (where Ras is localized) in response to activating signals. Thus, membrane recruitment is a necessary step (however, it may not be sufficient) for activation of RasGRFs.

Localization studies of RasGRF4 are important in determining the activation mechanisms of this protein. We have performed immunofluorescence localization studies in Hek 293T cells transiently transfected with RasGRF4, using RasGRF4 specific antibodies which we have raised. Our results show that RasGRF4 is primarily associated with the plasma membrane. RasGRF4 has a PDZ domain and a PDZ-binding motif. PDZ domains have been known to be important in targeting proteins to the plasma membrane. Therefore, the PDZ domain of RasGRF4 targets it to the plasma membrane by likely binding to transmembrane receptors or ion channels which harbour its binding sites. The PDZ-binding motif of RasGRF4 does not mediate membrane targeting. We used mutant constructs either lacking the PDZ domain or having the putative PDZ-binding motif deleted in immunofluorescence localization studies to characterize their roles in RasGRF4 localization. We also perform localization studies on cells which are treated with various stimuli such as growth factors, cNMP-elevating agents, intracellular calcium elevating agents and so on, in order to measure each stimuli's effects on the localization of RasGRF4.

Our previous results from the binding studies with RasGRF4-PDZ domain show an intramolecular interaction in RasGRF4 by the association of its PDZ domain and its own PDZ-binding motif. If such an intramolecular interaction in RasGRF4 is used to regulate its activity, then the mutant constructs, which either lacks the PDZ domain or has the mutated PDZ-binding motif, affects RasGRF4 activity.

Since RasGRF4 has a cNMP-binding domain it suggests that cNMP (preferably cAMP or cGMP) has regulatory roles on RasGRF4 activity and our recent work has indeed demonstrated activation of Ras/MAPK pathway by RasGRF4 in response to cAMP or cGMP analogues. We will perform cNMP binding assays to test for cNMP binding to this domain. Cyclic AMP is a secondary messenger for G-protein coupled receptors which activate adenylyl cyclases by coupling to G-proteins. Many of these G-coupled receptors have putative PDZ-binding motifs in their intracellular C-terminal ends which potentially bind to PDZ-containing proteins. Having both a PDZ domain and a putative cAMP-binding domain, RasGRF4 may be

involved in G-coupled receptor signaling pathways. We identify a receptor/receptors which bind specifically to the PDZ domain of RasGRF4 as binding leads to membrane targeting of RasGRF4 and to changes in RasGRF4 activity. We use several known G-coupled receptors such as beta-adrenergic receptors, NMDA receptor, Dopamine receptor and others. The later two are neuronal receptors and RasGRF4 was shown to be expressed strongly in the central nervous systems.

d) Determine the roles of Nedd4 in RasGRF4 regulation:

Since Nedd4 is a ubiquitin protein ligase, which we showed binds rasGRF4, it ubiquitinates and targets RasGRF4 for degradation. The mSOS2 as well as RasGRF2 were shown to be regulated by ubiquitination (46,47). We perform ubiquitination assays to measure RasGRF4 ubiquitination. Concurrently, stability studies (pulse-chase experiments) are also carried out to measure stability of RasGRF4.

In addition, since Nedd4 has a C2 domain which is a Ca2+-dependent lipid binding domain, we measure the effects of calcium on the localization and activity of RasGRF4.

MATERIALS AND METHODS

Identification of novel proteins interacting with Nedd4-WW domains

The method of identifying RasGRF4 is as follows. An expression library screen was used to identify proteins interacting with Nedd4-WW domains. GST-fusion proteins of individual WW domains of Nedd4 were constructed in pGEX2TK which contains a PKA phosphorylation site allowing radiolabeling of the fusion proteins with P32-ATP. The radiolabeled GST-fusion protein of Nedd4-WW2 domain was used as a probe to screen a 16-day mouse embryo expression library. About 106 cDNA clones were screened. A total of 17 independent positive clones were isolated and sequenced using dideoxy sequencing method. All isolated clones contained at least one PY motif and thus are biochemically true positives.

Among the positive clones isolated was Clone 7.7. Clone 7.7 is a novel protein, the partial amino sequence of which exhibits 75% identity and 95% similarity of that of the novel human brain cDNA called KIAA0313 isolated as part of Human Genome Project (Fig.2). Because of this remarkable high sequence similarity between them, we believe that Clone 7.7 is the mouse homologue of KIAA0313 and obtained the full-length cDNA of KIAA0313 from

the Kazusa DNA Research Institution which previously isolated this clone as part of Human Genome Project.

The present invention has been described in detail and with particular reference to the preferred embodiments; however, it will be understood by one having ordinary skill in the art that changes can be made without departing from the spirit and scope thereof. For example, where the application refers to proteins, it is clear that peptides and polypeptides may often be used. Likewise, where a gene is described in the application, it is clear that nucleic acid molecules or gene fragments may often be used.

All publications (including Genebank entries), patents and patent applications are incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety.

REFERENCES

1- Staub, O., et al. The EMBO Journal 15, 2371-2380 (1996).

2- Gawler,D., Zhang, L.-J., Reedijk, M., Tung, P.S. & Moran, M. Oncogene 10, 817-825 (1995).

3- Perin, M.S., Brose, N., Jahn, R. & Sudof, T.C. Journal of Biological Chemistry 266, 623-629 (1991).

4- Shirataki, H., et al. Molecular and Cellular Biology 13, 2061-2068 (1993).

5- Clark, J.D., et al. Cell 65, 1043-1051 (1991).

6- Ponting, C.P. & Parker, P.J. Protein Science 5, 162-166 (1996).

7- Plant, P., et al. Journal of Biological Chemistry 272, 32329 (1997).

8- Andre, B. & Springael, J.-Y. Biochemical and Biophysical Reasearch Communications 205, 1201-1205 (1994).

9- Bork, P. & Sudol, M. Trends in Biochemical Sciences 19, (1994).

10- Hofmann, K. & Bucher, P. FEBS Letters 358, 153-157 (1995).

11- Staub, O. & Rotin, D. Structure 4, 495-499 (1996).

12- Macias, M.M., et al. Nature 382, 646-649 (1996).

13- Ranganathan, R., Lu, K., Hunter, T. & Noel, J.P. Cell 89,875-886 (1997).

14- Chen, H.I. & Sudol, M. Proceedings of the National Academy of Science (USA) 92, 7819-7823 (1995).

15- Shimkets, R.a., et al. Cell 79, 407-414 (1994).

16- Hansson, J.H., et al. P.N.A.S. (USA) 92. 11495-11499 (1995).

17- Hansson, J.H., et al. Nature Genetics 11, 76-82 (1995).

18- Tamura, H., et al. Journal of Clinical Investigation 97, 1780-1784 (1996).

19- Staub, O., et al. The EMBO Journal 16, 6325-6336 (1997).

20- Jung, D., et al. Journal of Biological Chemistry 270, 27305-27310 (1995).

21- Broek, D., et al. Cell 48, 789 (1987).

22- Camus, C., et al. Oncogene 11, 951-959 (1995).

23- Boriack-Sjodin, P.A., et al. Nature 394, 337-343 (1998).

24- Fam, N.P., et al. Molecular and Cellular Biology 17, 1396-1406 (1997).

25- Pawson, T. Nature 373, 573-580 (1995).

26- Farnsworth, C.L., et al. Nature 376, 524-527 (1995).

27- Ebinu J.O., et al. Science 280, 1082-1086 (1998).

28- Christopher P. Ponting, et al. BioEssays 19, 469-479 (1997).

29- Stephen N. Gomperts. Cell 84. 659-662 (1996).

30- Chevesich, J., et al. Neuron 18, 95-105 (1997).

31- Songyang, Z., et al. Science 275, 73-77 (1997).

32- Huang, L., et al. Nat Struct Biol 4(8), 609-615 (1997).

33- Phillip G.N. Protein 14, 425-9 (1992).

34- Trier, M., et al. Cell 78, 787-798 (1994).

35- Taylor, S. & Shalloway, D. Current Biology 6, 1621 (1996).

36- Doskeland, S.O., et al. Methods Enzymology 159, 147-150 (1988).

37- Hock, B., et al. PNAS 95, 9779-9784 (1998).

38- Wu, J., et al. Science 262, 1065-1068 (1993).

39- Cook, S.J., et al. Science 262, 1069-1071 (1993).

40- Abriel, H., et al. (1998). J. Clin. Invest, In press.

41- Bouras, M., et al. Eur. J. Endocrinol. 139, 209-16 (1998).

42- Matias, G., et al. Virchows Arch 433, 103-111 (1998).

43- Kressmen, U., et al. Eur. J. Cancer 34, 737-744 (1998).

44- Abrams, S.I., et al. Semin. Oncology 23, 118-134 (1996).

45- Chan, David C., et al. The EMBO Journal 15, 1045-1054 (1996).

46- Nielsen, K.H., et al. Mol. Cell. Biol. 17, 7132-8 (1997).

47- – –

48- Weisskopf, Marc G., et al. Science, Vol. 265, 1878-1882 (1994).

49- Lyengar, Ravi. Science, Vol. 271, 461-463 (1996).

50- Shabb, John B., et al. The Journal of Biological Chemistry, Vol. 267, No. 9, 5723-5726 (199?).

51- Taylor, Susan S.. The Journal of Biological Chemistry, Vol. 264, No. 15, 8443-8446 (1989).

52- Frey, U., et al. Science, Vol. 260, 1661-1664 (1993).

53- Weber, Irene T.. Biochemistry, 26, 343-351 (1987).

54- Hammerschmidt, Matthias, et al. Genes & Development, 10:647-658 (1996).

55- Bailey, Craig H., et al. Proc Natl. Acad. Sci. USA, Vol. 93, 13445-13452 (1996).

56- Downward, J.. Nature, Vol. 396, 416-417 (1998).

57- Kawasaki, Kiroaki, et al. Science, Vol. 282, 2275-2279 (1998).

58- de Rool, Johan, et al. Nature, Vol. 396, 474-477 (1998).

59 Weber, et al. Biochemistry, 28:6/22-6127. (1989).

60 Weber, et al. Biochemistry, 26:343-351 (1987).

61 Su, et al. Science, 269:807-813 (1995).

We claim:

1. An isolated nucleic acid molecule encoding a polypeptide having RasGRF4 activity.

2. The nucleic acid molecule of claim 1, comprising all or part of the nucleic acid molecule of [SEQ ID NO:1].

3. An isolated nucleic molecule comprising at least 40% sequence identity to all or part of the nucleic acid molecule of [SEQ ID NO:1] , wherein the nucleic acid molecule encodes a polypeptide having RasGRF4 activity.

4. The molecule of any of claims 1 to 3 which is selected from a group consisting of mRNA, cDNA, sense DNA, anti-sense DNA, single-stranded DNA and double-stranded DNA.

5. A nucleic acid molecule encoding the amino acid sequence of [SEQ ID NO:2].

6. A nucleic acid molecule that encodes all or part of a RasGRF4 polypeptide or a polypeptide having RasGRF4 activity, wherein the sequence hybridizes to the nucleic acid molecule of all or part of [SEQ ID NO:1] under high stringency conditions.

7. The nucleic acid molecule of claim 6, wherein the high stringency conditions comprise a wash stringency of about 0.2X SSC, about 0.1% SDS, at about 50-65°C.

8. An isolated polypeptide having RasGRF4 activity and a CDC25 domain.

9. The polypeptide of claim 8, comprising all or part of the sequence of [SEQ ID NO:2].

10. An isolated polypeptide comprising at least 40% sequence identity to all or part of the polypeptide of [SEQ ID NO:2] , wherein the polypeptide has RasGRF4 activity.

11. A mimetic of the isolated polypeptide of any of claims 8 to 10, wherein the mimetic has RasGRF4 activity.

12. A recombinant nucleic acid molecule comprising a nucleic acid molecule of any of claim 1 to claim 7 and a promoter region, operatively linked so that the promoter enhances transcription of the nucleic acid molecule in a host cell.

13. A system for the expression of RasGRF4, comprising an expression vector and a nucleic acid molecule of any of claim 1 to claim 7 inserted in the expression vector.

14. The system of claim 13, wherein the expression vector comprises a plasmid or a virus.

15. A cell transformed by the expression vector of claim 14.

16. A method for expressing a polypeptide comprising: transforming an expression host with an expression vector including and culturing the expression host.

17. The method of claim 16, further comprising isolating the polypeptide.

18. The method of claim 16 or 17, wherein the expression host is selected from the group consisting of a plant, plant cell, bacterium, yeast, fungus, protozoa, algae, animal and animal cell.

19. A pharmaceutical composition, comprising all or part of the polypeptide or mimetic of any of claims 8 to 11, and a pharmaceutically acceptable carrier, auxiliary or excipient

20. A RasGRF4 specific antibody targeted to a region selected from the group consisting of the C-terminus, the CDC25 domain and the PDZ domain.

21. The peptide of claim 20, wherein the antibody is a monoclonal antibody or a polyclonal antibody.

22. A method of medical treatment of a disease, disorder or abnormal physical state, characterized by excessive RasGRF4 expression, concentration or activity, comprising administering a product that reduces or inhibits RasGRF4 polypeptide expression, concentration or activity.

23. The method of claim 22, wherein the product is an antisense nucleic acid molecule to all or part of the nucleic acid molecule of any of claims 1 to 7, the antisense nucleic acid molecule being sufficient to reduce or inhibit RasGRF4 polypeptide expression.

24. The method of claim 22, wherein the product comprises all or part of Nedd4.

25. The method of any of claims 22 to 24 wherein the disease, disorder or abnormal physical state is selected from a group consisting of cancer or diseases, disorders or abnormal physical states of the nervous system.

26. A method of medical treatment of a disease, disorder or abnormal physical state, characterized by inadequate RasGRF4 expression, concentration or activity, comprising administering a product that increases RasGRF4 polypeptide expression, concentration or activity.

27. The method of claim 26, wherein the product is a nucleic acid molecule comprising all or part of the nucleic acid molecule of any of claims 1 to 7, the DNA being sufficient to increase RasGRF4 polypeptide expression.

28. The method of claim 27, wherein the nucleic acid molecule is administered in a pharmaceutical composition comprising a carrier and a vector operably linked to the nucleic acid molecule.

Figure1

Rat Nedd4

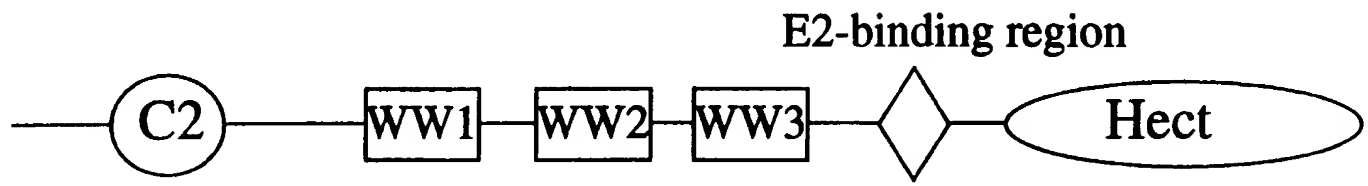


Figure 2: Clone 7.7 is the homolog of human clone KIAA0313

- Clone 7.7 exhibits 75% identity and 95% similarity of human clone KIAA0313.

```
Clone7.7      GGKDVSAEAESSSMVPVTTEEAKPVPMPAHIAVTPSTTKGLIARKEGRYREPPPTPPGYV
KIAA0313      GGKDVSIEAESSSLTSVTTEETKPVPMPAHIAVASSTTKGLIARKEGRYREPPPTPPGYI
              ****** ******:..*****:***********:.************************:

Clone7.7      GIPIADFPEGPCHPARKPPDYNVALQRSRMVARPTEAPAPG--QTPP-AAAASRPGSKPQ
KIAA0313      GIPITDFPEGHSHPARKPPDYNVALQRSRMVARSSDTAGPSSVQQPHGHPTSSRPVNKPQ
              ****:*****  .*********************.:::..*.   * *   .::*** .***

Clone7.7      WHKPSDADPRLAPFQAGFAGAEEDEDEQVSAV
KIAA0313      WHKPNESDPRLAPYQSQGFSTEEDEDEQVSAV
              ****.::******:*:    .:***********
```

Figure 3: **Schematic Diagram of RasGRF4**
(A)

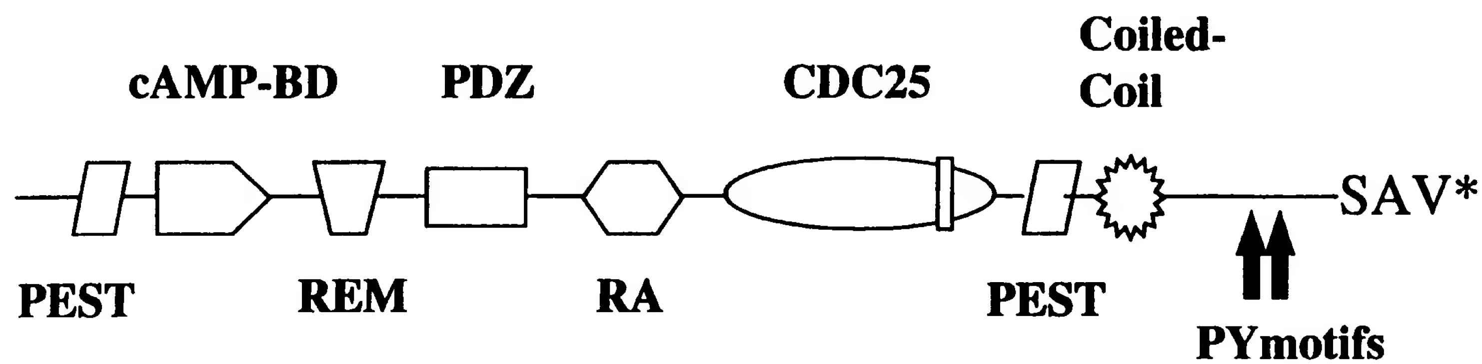


Figure 3B

KIAA0313 dna atg to stop -> 1-phase Translation

KIAA0313 = rasGRF4

```
DNA sequence     4618 b.p.    atgaaaccacta ... tttgcctccttc   linear

1/1                                         31/11
atg aaa cca cta gca atc cca gct aac cat gga gtt atg ggc cag cag gag aaa cac tca
M   K   P   L   A   I   P   A   N   H   G   V   M   G   Q   Q   E   K   H   S
61/21                                       91/31
ctt cct gca gat ttc aca aaa ctg cat ctt act gac agt ctc cac cca cag gtg acc cac
L   P   A   D   F   T   K   L   H   L   T   D   S   L   H   P   Q   V   T   H
121/41                                      151/51
gtt tct tct agc cat tca gga tgt agt atc act agt gat tct ggg agc agc agt ctt tct
V   S   S   S   H   S   G   C   S   I   T   S   D   S   G   S   S   S   L   S
181/61                                      211/71
gat atc tac cag gcc aca gaa agc gag gct ggt gat atg gac ctg agt ggg ttg cca gaa
D   I   Y   Q   A   T   E   S   E   A   G   D   M   D   L   S   G   L   P   E
241/81                                      271/91
aca gca gtg gat tcc gaa gac gac gac gat gaa gaa gac att gag aga gca tca gat cct
T   A   V   D   S   E   D   D   D   D   E   E   D   I   E   R   A   S   D   P
301/101                                     331/111
ctg atg agc agg gac att gtg aga gac tgc cta gag aag gac cca att gac cgg aca gat
L   M   S   R   D   I   V   R   D   C   L   E   K   D   P   I   D   R   T   D
361/121                                     391/131
gat gac att gaa caa ctc ttg gaa ttt atg cac cag ttg cct gct ttt gcc aat atg aca
D   D   I   E   Q   L   L   E   F   M   H   Q   L   P   A   F   A   N   M   T
421/141                                     451/151
atg tca gtg agg cga gaa ctc tgt gct gtg atg gtg ttc gca gtg gtg gaa aga gca ggg
M   S   V   R   R   E   L   C   A   V   M   V   F   A   V   V   E   R   A   G
481/161                                     511/171
acc ata gtg tta aat gat ggt gaa gag ctg gac tcc tgg tca gtg att ctc aat gga tct
T   I   V   L   N   D   G   E   E   L   D   S   W   S   V   I   L   N   G   S
541/181                                     571/191
gtg gaa gtg act tat cca gat gga aaa gca gaa ata ctg tgc atg gga aat agt ttt ggt
V   E   V   T   Y   P   D   G   K   A   E   I   L   C   M   G   N   S   F   G
601/201                                     631/211
gtc tct cct acc atg gac aaa gaa tac atg aaa gga gtg atg aga aca aag gtg gat gac
V   S   P   T   M   D   K   E   Y   M   K   G   V   M   R   T   K   V   D   D
661/221                                     691/231
tgc cag ttt gtc tgc ata gcc cag caa gat tac tgc cgt att ctc aat caa gta gaa aag
C   Q   F   V   C   I   A   Q   Q   D   Y   C   R   I   L   N   Q   V   E   K
721/241                                     751/251
aac atg caa aaa gtt gaa gag gaa gga gag att gtt atg gtg aaa gaa cac cga gaa ctt
N   M   Q   K   V   E   E   E   G   E   I   V   M   V   K   E   H   R   E   L
781/261                                     811/271
gat cga act gga aca aga aag gga cac att gtc atc aag ggt acc tca gaa agg tta aca
D   R   T   G   T   R   K   G   H   I   V   I   K   G   T   S   E   R   L   T
841/281                                     871/291
atg cat ttg gtg gaa gag cat tca gta gta gat cca aca ttc ata gaa gac ttt ctg ttg
M   H   L   V   E   E   H   S   V   V   D   P   T   F   I   E   D   F   L   L
901/301                                     931/311
acc tat agg act ttt ctt tct agc cca atg gaa gtg ggc aaa aag tta ttg gag tgg ttt
T   Y   R   T   F   L   S   S   P   M   E   V   G   K   K   L   L   E   W   F
961/321                                     991/331
aat gac ccg agc ctc agg gat aag gtt aca cgg gta gta tta ttg tgg gta aat aat cac
N   D   P   S   L   R   D   K   V   T   R   V   V   L   L   W   V   N   N   H
1021/341                                    1051/351
ttc aat gac ttt gaa gga gat cct gca atg act cga ttt tta gaa gaa ttt gaa aac aat
F   N   D   F   E   G   D   P   A   M   T   R   F   L   E   E   F   E   N   N
1081/361                                    1111/371
ctg gaa aga gag aaa atg ggt gga cac cta agg ctg ttg aat atc gcg tgt gct gct aaa
L   E   R   E   K   M   G   G   H   L   R   L   L   N   I   A   C   A   A   K
1141/381                                    1171/391
gca aaa aga aga ttg atg acg tta aca aaa cca tcc cga gaa gct cct ttg cct ttt atc
A   K   R   R   L   M   T   L   T   K   P   S   R   E   A   P   L   P   F   I
1201/401                                    1231/411
tta ctt gga ggc tct gag aag gga ttt gga atc ttt gtt gac agt gta gat tca ggt agc
L   L   G   G   S   E   K   G   F   G   I   F   V   D   S   V   D   S   G   S
```

cNMP-BD

REM

PDZ

1261/421 1291/431
aaa gca act gaa gca ggc ttg aaa cgg ggg gat cag ata tta gaa gta aat ggc caa aac
K A T E A G L K R G D Q I L E V N G Q N
1321/441 1351/451
ttt gaa aac att cag ctg tca aaa gct atg gaa att ctt aga aat aac aca cat tta tct
F E N I Q L S K A M E I L R N N T H L S
1381/461 1411/471
atc act gtg aaa acc aat tta ttt gta ttt aaa gaa ctt cta aca aga ttg tca gaa gag
I T V K T N L F V F K E L L T R L S E E
1441/481 1471/491
aaa aga aat ggt gcc ccc cac ctt cct aaa att ggt gac att aaa aag gcc agt cgc tac
K R N G A P H L P K I G D I K K A S R Y
1501/501 1531/511
tcc att cca gat ctt gct gta gat gta gaa cag gtg ata gga ctt gaa aaa gtg aac aaa
S I P D L A V D V E Q V I G L E K V N K
1561/521 1591/531
aaa agt aaa gcc aac act gtg gga gga agg aac aag ctg aaa aag ata ctc gac aag act
K S K A N T V G G R N K L K K I L D K T
1621/541 1651/551
cgg atc agt atc ttg cca cag aaa cca tac aat gat att ggg att ggt cag tct caa gat
R I S I L P Q K P Y N D I G I G Q S Q D
1681/561 1711/571
gac agc ata gta gga tta agg cag aca aag cac atc cca act gca ttg cct gtc agt gga
D S I V G L R Q T K H I P T A L P V S G
1741/581 1771/591
acc tta tca tcc agt aat cct gat tta ttg cag tca cat cat cgc att tta gac ttc agt
T L S S N P D L L Q S H H R I L D F S
1801/601 1831/611
gct act cct gac ttg cca gat caa gtg cta agg gtt ttt aag gct gat cag caa agc cgc
A T P D L P D Q V L R V F K A D Q Q S R
1861/621 1891/631
tac atc atg atc agt aag gac act aca gca aag gaa gtg gtc att cag gct atc agg gag
Y I M I S K D T T A K E V V I Q A I R E
1921/641 1951/651
ttt gct gtt act gcc acc ccg gat caa tat tca cta tgt gag gtc tct gtc aca cct gag
F A V T A T P D Q Y S L C E V S V T P E
1981/661 2011/671
gga gta atc aaa caa aga aga ctt cca gat cag ctt tcc aaa ctt gca gac aga ata caa
G V I K Q R R L P D Q L S K L A D R I Q
2041/681 2071/691
ctg agt gga agg tat tat ctg aaa aac aac atg gaa aca gaa act ctt tgt tca gat gaa
L S G R Y Y L K N N M E T E T L C S D E
2101/701 2131/711
gat gct cag gag ttg ttg aga gag agt caa att tcc ctc ctt cag ctc agc act gtg gaa
D A Q E L L R E S Q I S L L Q L S T V E
2161/721 2191/731
gtt gca aca cag ctc tct atg cga aat ttt gaa ctc ttt cgc aac att gaa cct act gaa
V A T Q L S M R N F E L F R N I E P T E
2221/741 2251/751
tat ata gat gat tta ttt aaa ctc aga tca aaa acc agc tgt gcc aac ctg aag aga ttt
Y I D D L F K L R S K T S C A N L K R F
2281/761 2311/771
gaa gaa gtc att aac cag gaa aca ttt tgg gta gca tct gaa att ctc aga gaa aca aac
E E V I N Q E T F W V A S E I L R E T N
2341/781 2371/791
cag ctg aag agg atg aag atc att aag cat ttc atc aag ata gca ctg cac tgt agg gaa
Q L K R M K I I K H F I K I A L H C R E
2401/801 2431/811
tgc aag aat ttt aac tca atg ttt gca atc atc agt ggc cta aac ctg gca cca gtg gca
C K N F N S M F A I I S G L N L A P V A
2461/821 2491/831
aga ctg cga acg acc tgg gag aaa ctt ccc aat aaa tac gaa aaa cta ttt caa gat ctc
R L R T T W E K L P N K Y E K L F Q D L
2521/841 2551/851
caa gac ctg ttt gat cct tcc aga aac atg gca aaa tat cgt aat gtt ctc aat agt caa
Q D L F D P S R N M A K Y R N V L N S Q
2581/861 2611/871
aat cta caa cct ccc ata atc cct cta ttc cca gtt atc aaa aag gat ctc acc ttc ctt
N L Q P P I I P L F P V I K K D L T F L

RA

CDC25

```
2641/881                                  2671/891
cac gaa gga aat gac tca aaa gta gac ggg ctg gtc aat ttt gag aag cta agg atg att
H   E   G   N   D   S   K   V   D   G   L   V   N   F   E   K   L   R   M   I
2701/901                                  2731/911
gca aaa gaa att cgt cac gtt ggc cga atg gct tca gtg aac atg gac cct gcc ctc atg
A   K   E   I   R   H   V   G   R   M   A   S   V   N   M   D   P   A   L   M
2761/921                                  2791/931
ttc agg act cgg aag aag aaa tgg cgg agt ttg ggg tct ctc agc cag ggt agt aca aat
F   R   T   R   K   K   K   W   R   S   L   G   S   L   S   Q   G   S   T   N
2821/941                                  2851/951
gca aca gtg cta gat gtt gct cag aca ggt ggt cat aaa aag cgg gta cgt cgt agt tcc
A   T   V   L   D   V   A   Q   T   G   G   H   K   K   R   V   R   R   S   S
2881/961                                  2911/971
ttt ctc aat gcc aaa aag ctt tat gaa gat gcc caa atg gct cga aaa gtg aag cag tac
F   L   N   A   K   K   L   Y   E   D   A   Q   M   A   R   K   V   K   Q   Y
2941/981                                  2971/991
ctt tcc aat ttg gag cta gaa atg gac gag gag agt ctt cag aca tta tct ctg cag tgt
L   S   N   L   E   L   E   M   D   E   E   S   L   Q   T   L   S   L   Q   C
3001/1001                                 3031/1011
gag cca gca acc aac aca ttg cct aag aat cct ggt gac aaa aag cct gtc aaa tcc gag
E   P   A   T   N   T   L   P   K   N   P   G   D   K   K   P   V   K   S   E
3061/1021                                 3091/1031
acc tct cca gta gct cca agg gca ggg tca caa cag aaa gct cag tcc ctg cca cag ccc
T   S   P   V   A   P   R   A   G   S   Q   Q   K   A   Q   S   L   P   Q   P
3121/1041                                 3151/1051
cag cag cag cca cca cca gca cat aaa atc aac cag gga cta cag gtt ccc gcc gtg tcc
Q   Q   Q   P   P   P   A   H   K   I   N   Q   G   L   Q   V   P   A   V   S
3181/1061                                 3211/1071
ctt tat cct tca cgg aag aaa gtg ccc gta aag gat ctc cca cct ttt ggc ata aac tct
L   Y   P   S   R   K   K   V   P   V   K   D   L   P   P   F   G   I   N   S
3241/1081                                 3271/1091
cca caa gct tta aaa aaa att ctt tct ttg tct gaa gaa gga agt ttg gaa cgt cac aag
P   Q   A   L   K   K   I   L   S   L   S   E   E   G   S   L   E   R   H   K
3301/1101                                 3331/1111
aaa cag gct gaa gat aca ata tca aat gca tct tcg cag ctt tct tct cct cct act tct
K   Q   A   E   D   T   I   S   N   A   S   S   Q   L   S   S   P   P   T   S
3361/1121                                 3391/1131
cca cag agt tct cca agg aaa ggc tat act ttg gct ccc agt ggt act gtg gat aat ttt
P   Q   S   S   P   R   K   G   Y   T   L   A   P   S   G   T   V   D   N   F
3421/1141                                 3451/1151
tca gat tct ggt cac agt gaa att tct tca cga tcc agt att gtt agc aat tcg tct ttt
S   D   S   G   H   S   E   I   S   S   R   S   S   I   V   S   N   S   S   F
3481/1161                                 3511/1171
gac tca gtg cca gtc tca ctg cac gat gag agg cgc cag agg cat tct gtc agc atc gtg
D   S   V   P   V   S   L   H   D   E   R   R   Q   R   H   S   V   S   I   V
3541/1181                                 3571/1191
gaa aca aac cta ggg atg ggc agg atg gag agg cgg acc atg att gaa cct gat cag tat
E   T   N   L   G   M   G   R   M   E   R   R   T   M   I   E   P   D   Q   Y
3601/1201                                 3631/1211
agc ttg ggg tcc tat gca cca atg tcc gag ggc cga ggc tta tat gct aca gct aca gta
S   L   G   S   Y   A   P   M   S   E   G   R   G   L   Y   A   T   A   T   V
3661/1221                                 3691/1231
att tct tct cca agc aca gag gaa ctt tcc cag gat cag ggg gat cgc gcg tca ctt gat
I   S   S   P   S   T   E   E   L   S   Q   D   Q   G   D   R   A   S   L   D
3721/1241                                 3751/1251
gct gct gac agt ggc cgt ggg agc tgg acg tca tgc tca agt ggc tcc cat gat aat ata
A   A   D   S   G   R   G   S   W   T   S   C   S   S   G   S   H   D   N   I
3781/1261                                 3811/1271
cag acg atc cag cac cag aga agc tgg gag act ctt cca ttc ggg cat act cac ttt gat
Q   T   I   Q   H   Q   R   S   W   E   T   L   P   F   G   H   T   H   F   D
3841/1281                                 3871/1291
tat tca ggg gat cct gca ggt tta tgg gca tca agc agc cat atg gac caa att atg ttt
Y   S   G   D   P   A   G   L   W   A   S   S   S   H   M   D   Q   I   M   F
3901/1301                                 3931/1311
tct gat cat agc aca aag tat aac agg caa aat caa agt aga gag agc ctt gaa caa gcc
S   D   H   S   T   K   Y   N   R   Q   N   Q   S   R   E   S   L   E   Q   A
3961/1321                                 3991/1331
cag tcc cga gca agc tgg gcg tct tcc aca ggt tac tgg gga gaa gac tca gaa ggt gac
Q   S   R   A   S   W   A   S   S   T   G   Y   W   G   E   D   S   E   G   D
```

Insertion unique to rasGRF

Continuation of the CDC25 domain

4021/1341 4051/1351
aca ggc aca ata aag cgg agg ggt gga aag gat gtt tcc att gaa gcc gaa agc agt agc
T G T I K R R G G K D V S I E A E S S S

4081/1361 4111/1371
cta acg tct gtg act acg gaa gaa acc aag cct gtc ccc atg cct gcc cac ata gct gtg
L T S V T T E E T K P V P M P A H I A V

4141/1381 4171/1391
gca tca agt act aca aag ggg ctc att gca cga aag gag ggc agg tat cga gag ccc ccg
A S S T T K G L I A R K E G R Y R E P P

4201/1401 4231/1411
ccc acc cct ccc ggc tac att gga att ccc att act gac ttt cca gaa ggg cac tcc cat
P T P P G Y I G I P I T D F P E G H S H

4261/1421 4291/1431 PY motifs
cca gcc agg aaa ccg ccg gac tac aac gtg gcc ctt cag aga tcg cgg atg gtc gca cga
P A R K P P D Y N V A L Q R S R M V A R

4321/1441 4351/1451
tcc tcc gac aca gct ggg cct tca tcc gta cag cag cca cat ggg cat ccc acc agc agc
S S D T A G P S S V Q Q P H G H P T S S

4381/1461 4411/1471
agg cct gtg aac aaa cct cag tgg cat aaa ccg aac gag tct gac ccg cgc ctc gcc cct
R P V N K P Q W H K P N E S D P R L A P

4441/1481 4471/1491
tat cag tcc caa ggg ttt tcc acc gag gag gat gaa gat gaa caa gtt tct gct gtt tga
Y Q S Q G F S T E E D E D E Q V S A V * PDZ binding motif

4501/1501 4531/1511
ggc aca gac ttt tct gga agc aga gcg agc cac ctg aaa gga gag cac aag aag acg tcc
G T D F S G S R A S H L K G E H K K T S

4561/1521 4591/1531
tga gca ttg gag cct tgg aac tca cat tct gag gac ggt gga cca gtt tgc ctc ctt c
* A L E P W N S H S E D G G P V C L L

Figure 4A: Alignment of CDC25 domains from several RasGEF/GRF including RasGRF4

```
                   SCR1
CDC25_yeast_       LLDIDSLDYAKQLTIKEHSLFYKISPFECLDRTWGNKYCN-MGGSKNITEFISNSNHLTN
SCD25_yeast_       ILAVDPVLFATQLTILEHEIYCEITIFDCLQKIWKNKYTKSYGASPGLNEFISFANKLTN
RasGEF_aimless_    IYDIDEEEIARQLTLIEFEIYRNIKPPELLNQSWNKTKLK--SRAPNVLKMIDRFNSVSM
RasGRF2_mouse_     FETLSAMELAEQITLLDHIVFRSIPYEEFLGQGWMKLDKN--ERTPYIMKTSQHFNEMSN
RasGRF4            LLQLSTVEVATQLSMRNFELFRNIEPTEYIDDLFKLRSKT---SCANLKRFEEVINQETF
SOS_human_         LLTLHPIEIARQLTLLESDLYRAVQPSELVGSVWTKEDKE--INSPNLLKMIRHTTNLTL
                   :  :       * *::: :  :: :  : :   :                :  .    .  :

CDC25_yeast_       YVSFMIVKQTDIKKRIQLIQFFINVAAHCHELNNFSSLTAIISALYSSPIYRLKRTWAAV
SCD25_yeast_       FISYSVVKEADKSKRAKLLSHFIFIAEYCRKFNNFSSMTAIISALYSSPIYRLEKTWQAV
RasGEF_aimless_    WVATMIIQTTKVKARARMMTRFIKIADHLKNLNNYNSLMAIIAGLNFSSVYRLKYTREEL   SCR2
RasGRF2_mouse_     LVASQIMNYADISSRPNAIEKWVAVADICRCLHNYNGVLEITSALNRSPIYRLKKTWAKV
RasGRF4            WVASEILRETNQLKRMKIIKHFIKIALHCRECKNFNSMFAIISGLNLAPVARLRTTWEKL
SOS_human_         WFEKCIVETENLEERVAVVSRIIEILQVFQELNNFNGVLEVVSAMNSSPVYRLDHTFEQI
                    .    ::.  .  *   :   : :    :   :*:..:  :  :.:   :.:  **   *   :

CDC25_yeast_       PEEYKKLLEELNTLMDSAKNFIRYRQLLKSIGD-FPCVPFFGVYLSDLTFTANGNPDFL-
SCD25_yeast_       IPQTRDLLQSLNKLMDPKKNFINYRNELKSLHS-APCVPFFGVYLSDLTFTDSGNPDYLV   SCR3
RasGEF_aimless_    SAQTMRTYSDLEKIMNSEGSFKTYRTRLQNVP---PMLPYLGVHLTDLTFIDE-NPNNFV
RasGRF2_mouse_     SKQTKALMDKLQKTVSSEGRFKNLRETLKNCN--PPAVPYLGMYLTDLAFIEEGTPNFT-
RasGRF4            PNKYEKLFQDLQDLFDPSRNMAKYRNVLNSQNLQPPIIPLFPVIKKDLTFLHEGNDSKV-
SOS_human_         PSRQKKILEEAHELS--EDHYKKYLAKLRSIN--PPCVPFFGIYLTNILKTEEGNPEVLK
                    .       .. .                 *..  * :*  :  :   .::    .  .  .

CDC25_yeast_       --------HRNTVLVNFGKRVRILEILKEISV
SCD25_yeast_       LEHGLKGVHDEKKYINFNKRSRLVDILQEIIY
RasGEF_aimless_    TDVG----GKQVSLINFTKRTLVFKIISLIQE
RasGRF2_mouse_     ----------EEGLVNFSKMRMISHIIREIRQ
RasGRF4            -----------DGLVNFEKLRMIAKEIRHVGR
SOS_human_         R--------HGKELINFSKRRKVAEITGEIQQ
                            :** *   :  .    :
```

Boxes = Structurally Conserved Region (SCR) 1, 2 & 3

Figure 4B

KIAA0313 = rasGRF4
GRF2 = rasGRF2

CDC25 Domain: KIAA0313 vs GRF2

CLUSTAL W (1.74) multiple sequence alignment

```
RasGRF2_mouse_      ATANILRALSQDDQD-DIHLKLEDIIQMTDCPKAECFETLSAMELAEQITLLDHIVFRSI
KIAA0313            RIQLSGRYYLKNNMETETLCSDEDAQELLRESQISLLQ-LSTVEVATQLSMRNFELFRNI
                          *   ::: :  :    . **  ::    .: . :: **::*:* *::: :. :**.*

RasGRF2_mouse_      PYEEFLGQGWMKLDKNERTPYIMKTSQHFNEMSNLVASQIMNYADISSRPNAIEKWVAVA
KIAA0313            EPTEYIDDLFKLRSKTSCAN-LKRFEEVINQETFWVASEILRETNQLKRMKIIKHFIKIA
                       *::.: :   .*.. :  : : .: :*: :  ***:*:. ::  .* : *:::: :*

RasGRF2_mouse_      DICRCLHNYNGVLEITSALNRSPIYRLKKTWAKVSKQTKALMDKLQKTVSSEGRFKNLRE
KIAA0313            LHCRECKNFNSMFAIISGLNLAPVARLRTTWEKLPNKYEKLFQDLQDLFDPSRNMAKYRN
                     **  :*:*.:: * *.** :*: **:.** *:.:: : *::.**. .... .: : *:

RasGRF2_mouse_      TLK--NCNPPAVPYLGMYLTDLAFIEEGTPNFTEEGLVNFSKMRMIS----HIIR-----
KIAA0313            VLNSQNLQPPIIPLFPVIKKDLTFLHEGN-DSKVDGLVNFEKLRMIAKEIRHVGRMASVN
                    .*:  * :** :* : :  .**:*:.**. : . :*****.*:***:    *: *

RasGRF2_mouse_      -------EIR-----------Q---------FQQT---AYRI-------------DQQ--
KIAA0313            MDPALMFRTRKKKWRSLGSLSQGSTNATVLDVAQTGGHKKRVRRSSFLNAKKLYEDAQMA
                           . *           *         . **    *:              * *

RasGRF2_mouse_      PKVIQYLLDKALVIDEDSLYELSLKIEPRLPA
KIAA0313            RKVKQYLSNLELEMDEESLQTLSLQCEPATNT
                     ** *** :  * :**:**  ***: **   :
```

Insertion unique to rasGRF4

Figure 5: RasGRF4-REM domain

```
CDC25             -IRGGTKEALIEHLT-SHELVDAAFNVTMLITFRSILT-TREFFYALIYRY-
Sos_mouse_        -IKGGTVVKLIERLT-YHMYADPNF-VRTFLTYRSFCK-PQELLNLLIERFE
RasGEF_aimless_   VVKFASLNKLVEHLT-HDSKHDLQFLKTFLMTYQSFCT-PEKLMSKLQQRY-
GRF2_mouse_       -IRYASVEALLERLT-DLRFLSIDFLNTFLHTYRIFTT-ATVVLAKLSDIY-
RasGRF4           -IKG-TSERLTMHLVEEHSVVDPTFIEDFLLTYRTFLSSPMEVGKKLLEWFN
                   ::  :   *  :*.      .  *    : *:: :  . .  .   *   :
```

Figure 6A:

Overall structure comparison between RasGRF4 and other known mammalian RasGEF/RasGRF

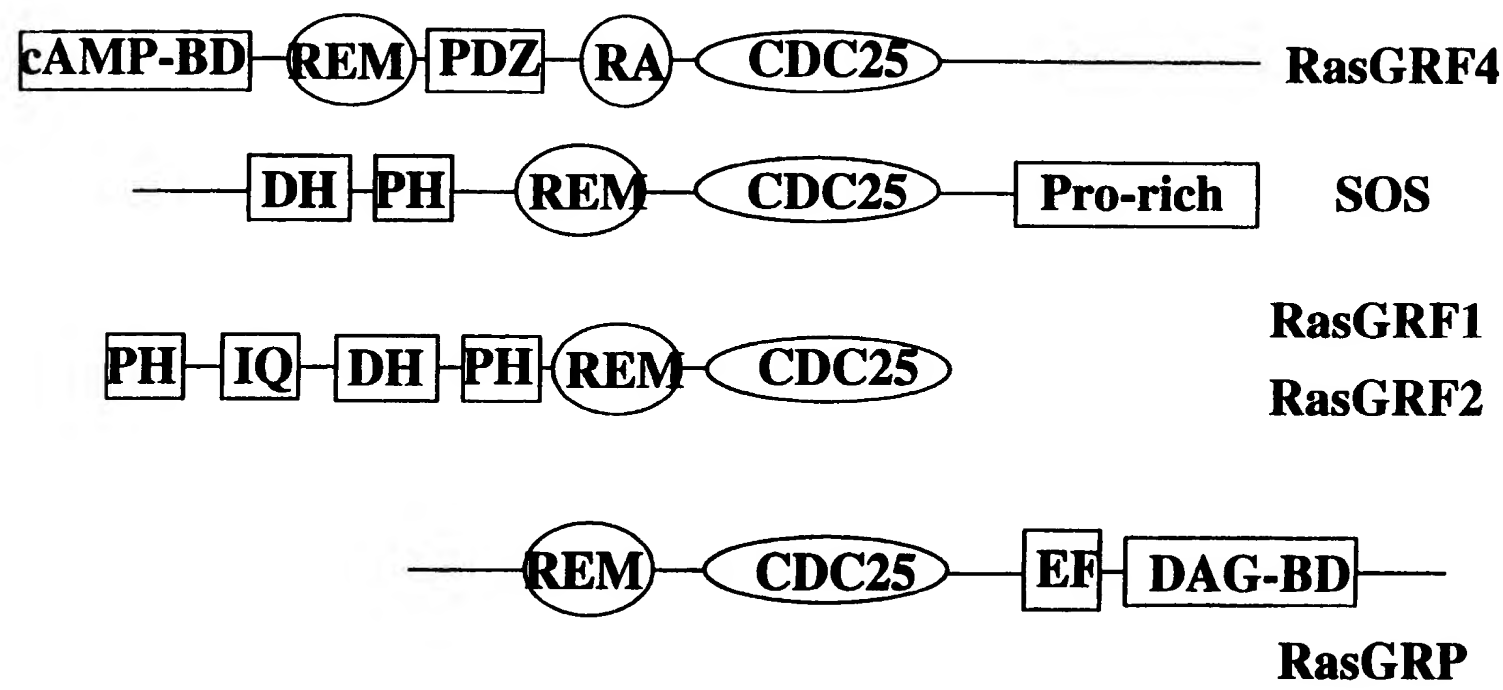


Fig.6B

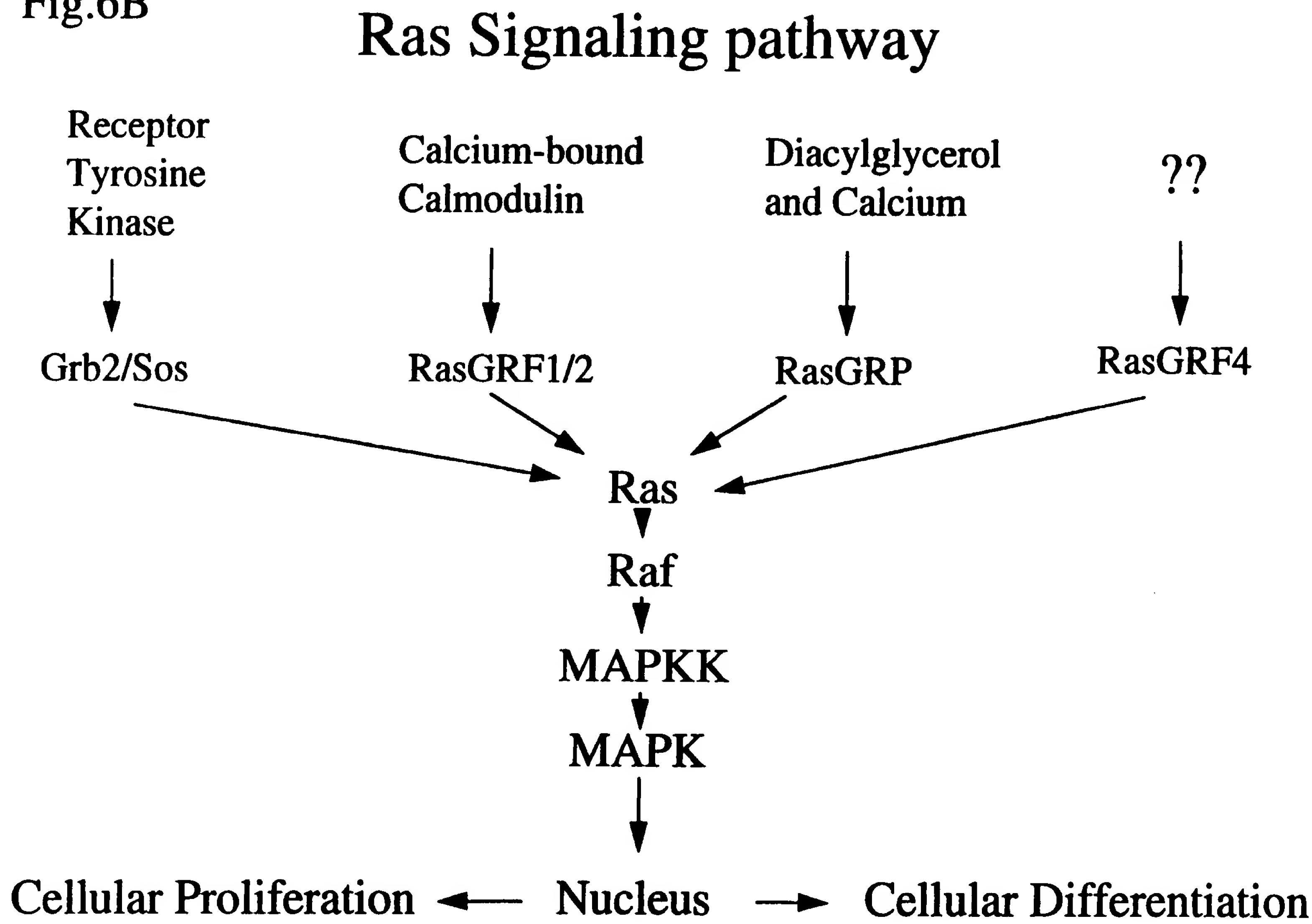

Ras Signaling pathway
Receptor Tyrosine Kinase
Calcium-bound Calmodulin
Diacylglycerol and Calcium
??
Grb2/Sos
RasGRF1/2
RasGRP
RasGRF4
Ras
Raf
MAPKK
MAPK
Cellular Proliferation
Nucleus
Cellular Differentiation

Figure 7: RasGRF4-PDZ domain

```
PSD-95_Mouse_      ---------------------LGFNIVGG----EDGEGIFISFILAGGPADLSGELRKGD
PTP-BAS_human_     -----REITLVNLK-KDAKYGLGFQIIGGEKMGRLDLGIFISSVAPGGPADLDGCLKPGD
LIN7_C.            AAEGHAHPRIVEL--PKTDQGLGFNVMGGK---EQNSPIYISRIIPGGVADRHGGLKRGD
RasGRF4            ACAAKAKRRLMTLTKPSREAPLPFILLGGS---EKGFGIFVDSVDSGSKATE-AGLKRGD
                                        * * ::**    . .  *::. : .*. *   . *: **

PSD-95_Mouse_      QILSVNGVDLRNASHEQAAIALKN
PTP-BAS_human_     RLISVNSVSLEGVSHHAAIEILQN
LIN7_C.            QLIAVNG-NVEAECHEKAVDLLKS
RasGRF4            QILEVNGQNFENIQLSKAMEILRN
                   ::: **. ...      *   *:.
```

Figure 8: RasGRF4-cAMP binding domain

```
RasGRF4      AFANMTMSVRRELCAVMVFAVVERAGTIVLNDGEELDSWSVILNGSVEVTYPDGKAEILC
T20G5.5      ALSHLSTMVKRQLSNFVKVEQYVHAGSVVFRQGEIGVYWYIVLKGAVEVNVNGKIVCLLR
PRKAR1B      LFAHLDDNE-RSDIFDAMFPVTHIAGETVIQQGNEGDNFYVVDQGEVDVYVNGEWVTNIS
              ::::    *.     .     **  *:.:*:     : :: :* *:*   .  .  :

RasGRF4      MGNSFGVSPTMDKEYMKGVMRTKVDDCQFVCIAQQDYCRILNQVEKNMQKVEEEGEIVM
T20G5.5      EGDDFGKLALVNDLPRAATIVTYEDDSMFLVVDKHHFNQILHQVEANTVRLKDYGEDVL
PRKAR1B      EGGSFGELALIYGTPRAATVKAKT-DLKLWGIDRDSYRRILMGSTLRKRKMYEE-----
                *..**  . :     ..: :   *  :  : :. : :**     .  :: :
```

Figure 9: RasGRF4-RA domain

```
dgk-1a_ce_     -----REDFE---IIRVFDGNNS---YRSQIS--------RNIVVAKHVSVQQVRDAALR
RasGRF4        --HHRILDFS---ATPDLPDQVLRVFKADQQS--------RYIMISKDTTAKEVVIQAIR
RalGDS_h_      SILVTSQDKAPSVISRVLKKNNRDSAVASEYELVQLLPGERELTIPASANVFYAMDGASH
                      *             :  :        .: .       * : :.  ...  .   * :

dgk-1a_ce_     RFHI--NDTPERYYIT-QVVGEVEEEILED----PVP------LRNVKRPEGKRAQIFIR
RasGRF4        EFAV--TATPDQYSLC-EVSVTPEGVIKQR----RLP------DQLSK--LADRIQLSGR
RalGDS_h_      DFLLRHGEGPLLLHLASPVARLPQELLRVREEGAPFPGSRPQGGRLHGHCSEEEAPLAYR
               * :     *    :   *     :  :       .*          :      .. :  *

dgk-1a_ce_     YYD------
RasGRF4        YYLKNNME-
RalGDS_h_      SHGVHTRCG
                :
```

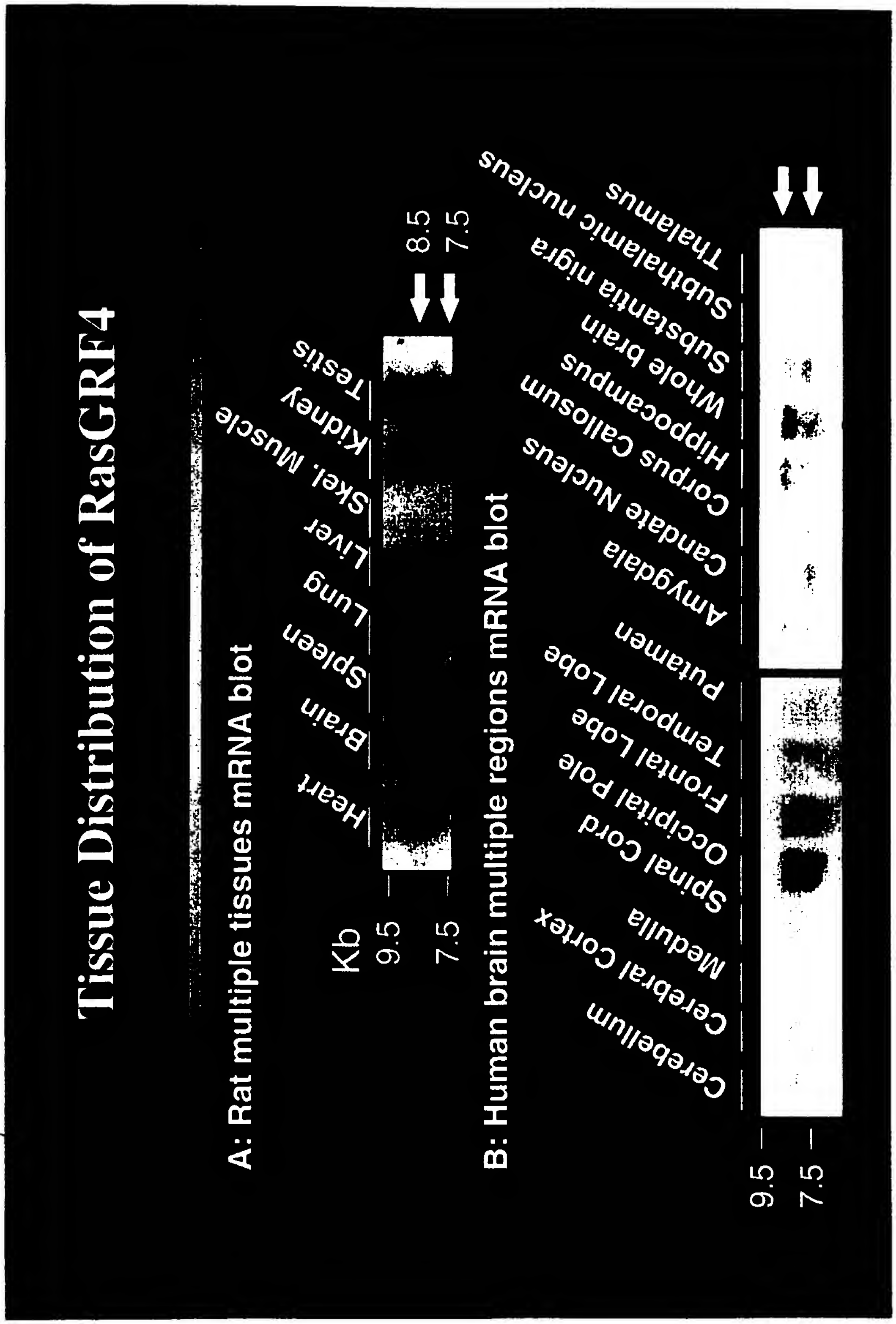
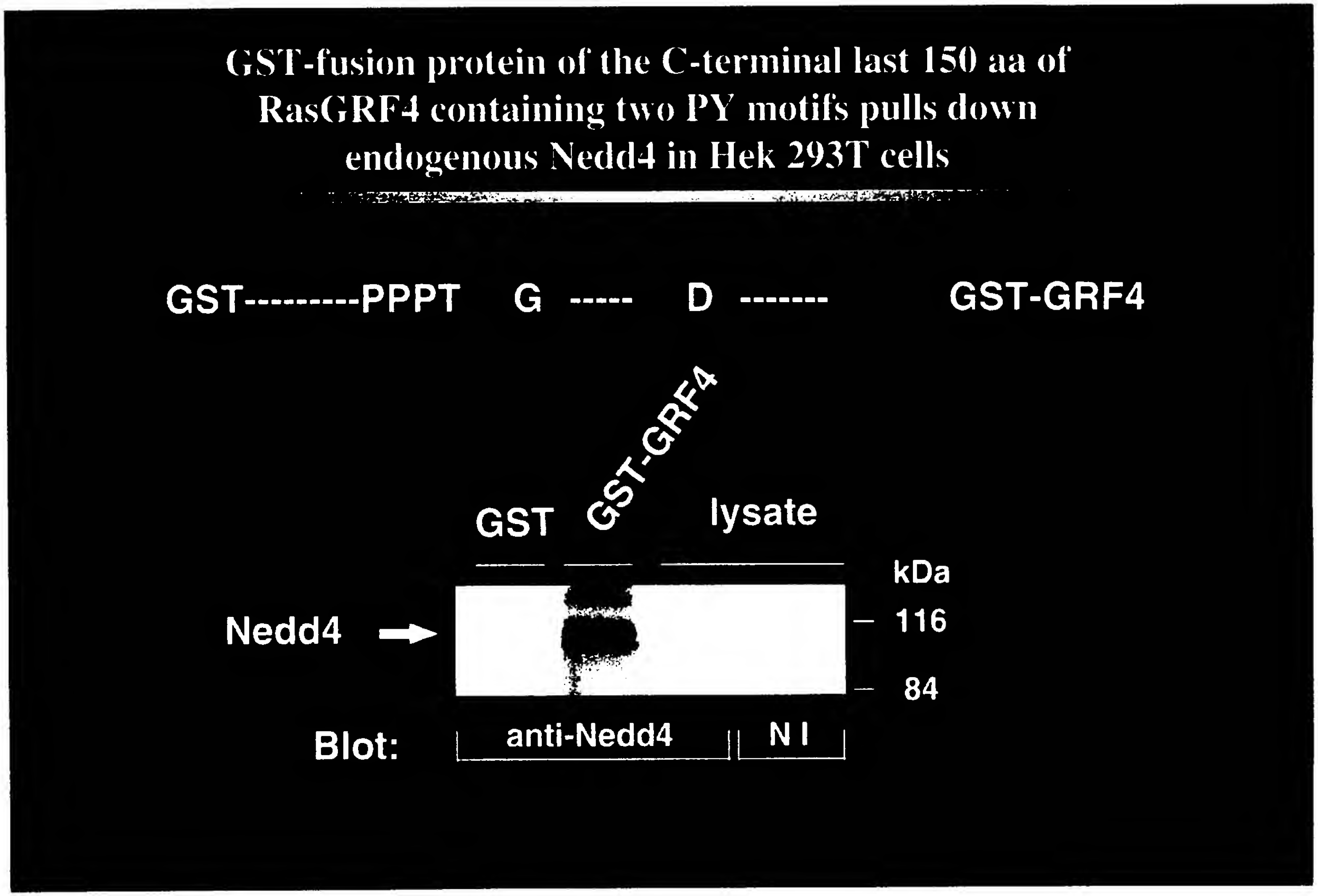

Tissue Distribution of RasGRF4
A: Rat multiple tissues mRNA blot
Kb
9.5
7.5
Heart
Brain
Spleen
Lung
Liver
Skel. Muscle
Kidney
Testis
8.5
7.5
B: Human brain multiple regions mRNA blot
Cerebellum
Cerebral Cortex
Medulla
Spinal Cord
Occipital Pole
Frontal Lobe
Temporal Lobe
Putamen
Amygdala
Candate Nucleus
Corpus Callosum
Hippocampus
Whole brain
Substantia nigra
Subthalamic nucleus
Thalamus
9.5
7.5

Fig. 10

Fig. 11


GST-fusion protein of the C-terminal last 150 aa of RasGRF4 containing two PY motifs pulls down endogenous Nedd4 in Hek 293T cells
GST---------PPPT G ----- D -------
GST-GRF4
GST
GST-GRF4
lysate
kDa
116
84
Nedd4
Blot:
anti-Nedd4
N I

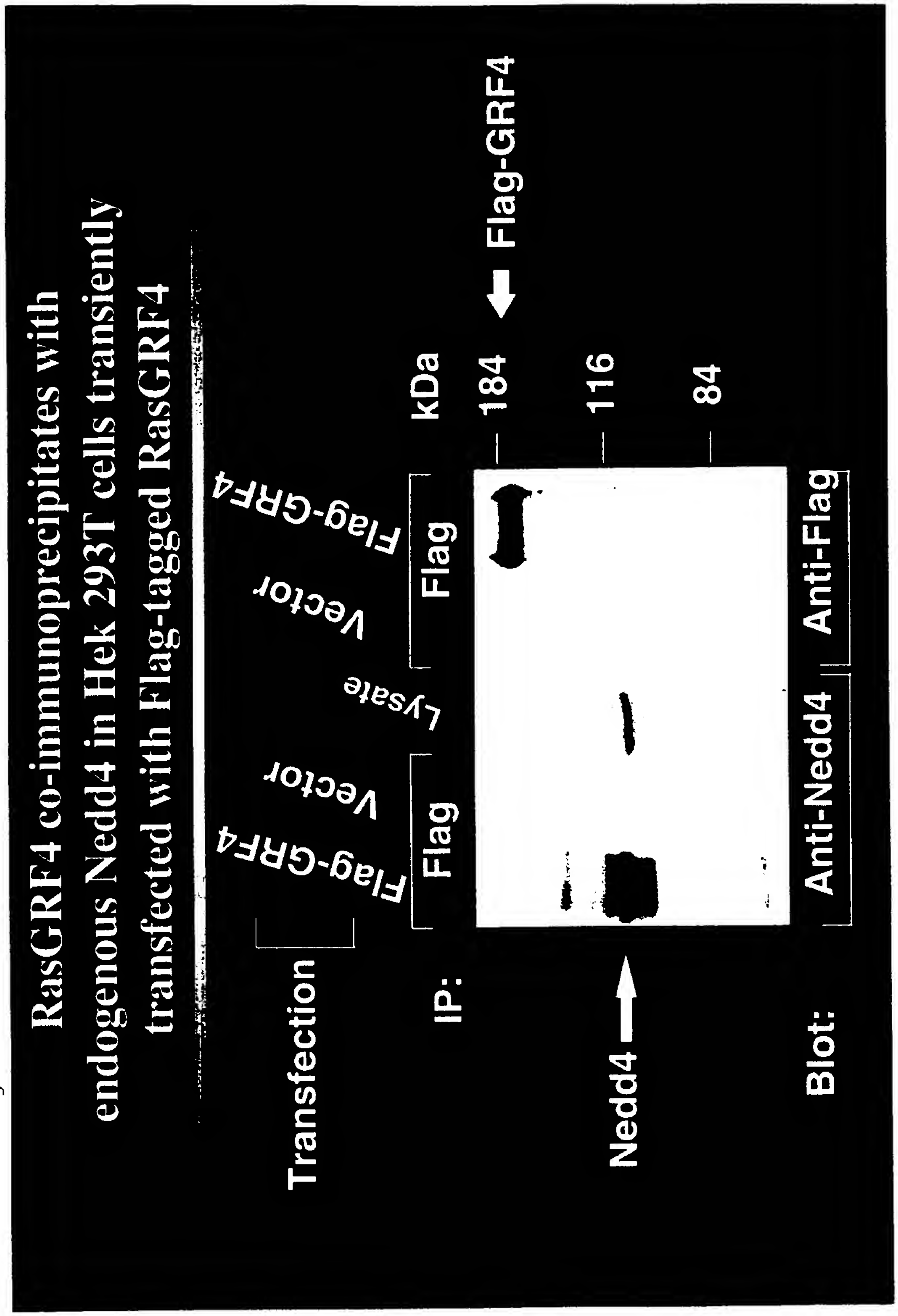

RasGRF4 co-immunoprecipitates with endogenous Nedd4 in Hek 293T cells transiently transfected with Flag-tagged RasGRF4
Transfection
Flag-GRF4
Vector
Lysate
Vector
Flag-GRF4
IP:
Flag
Flag
kDa
184
116
84
Flag-GRF4
Nedd4
Blot:
Anti-Nedd4
Anti-Flag

Fig. 12

Fig.13

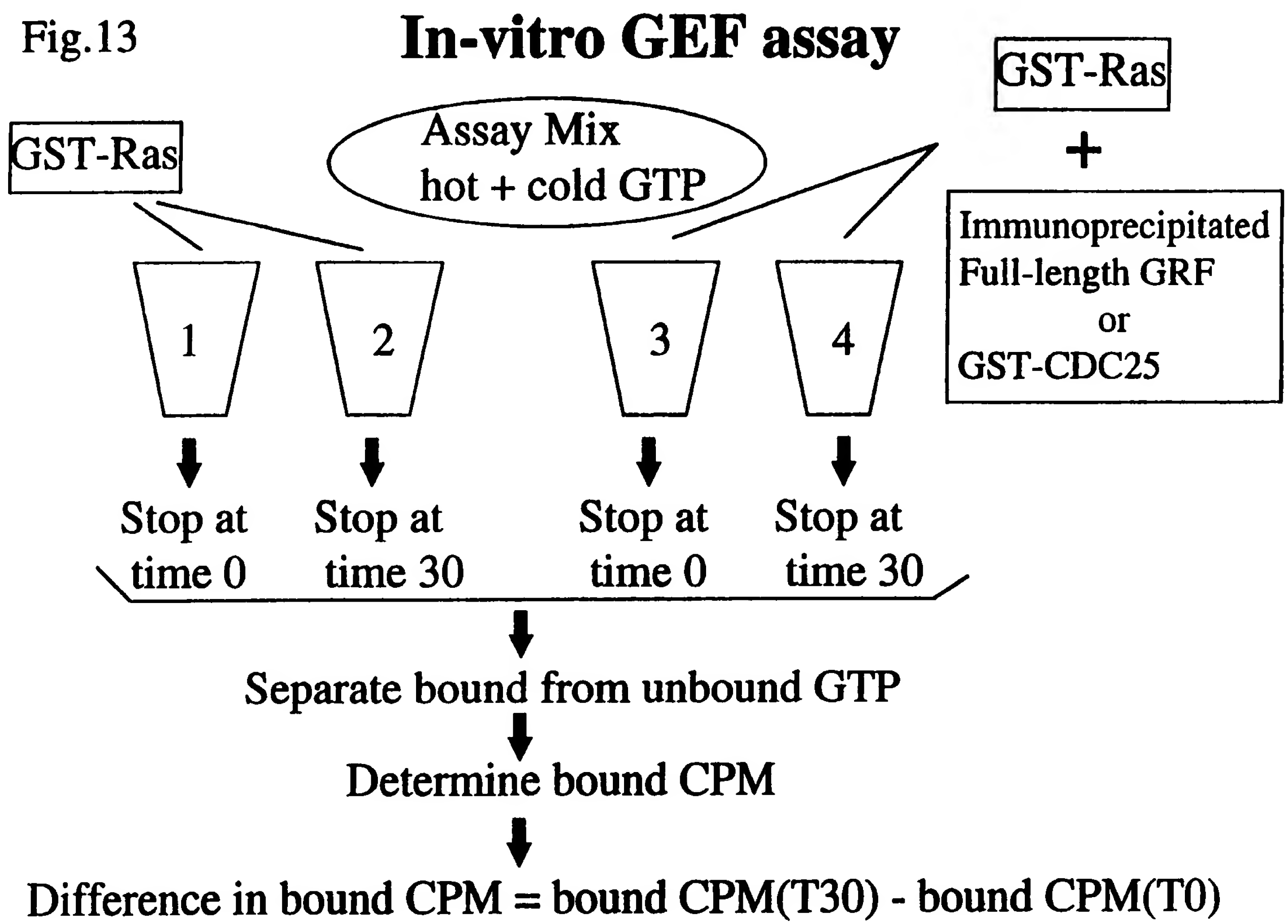

In-vitro GEF assay
GST-Ras
Assay Mix
hot + cold GTP
GST-Ras
+
Immunoprecipitated
Full-length GRF
or
GST-CDC25
1
2
3
4
Stop at
time 0
Stop at
time 30
Stop at
time 0
Stop at
time 30
Separate bound from unbound GTP
Determine bound CPM
Difference in bound CPM = bound CPM(T30) - bound CPM(T0)

Fig.14 RasGRF4 is active on Ras : in-vitro GEF assay using immunoprecipitated full-length RasGRF4

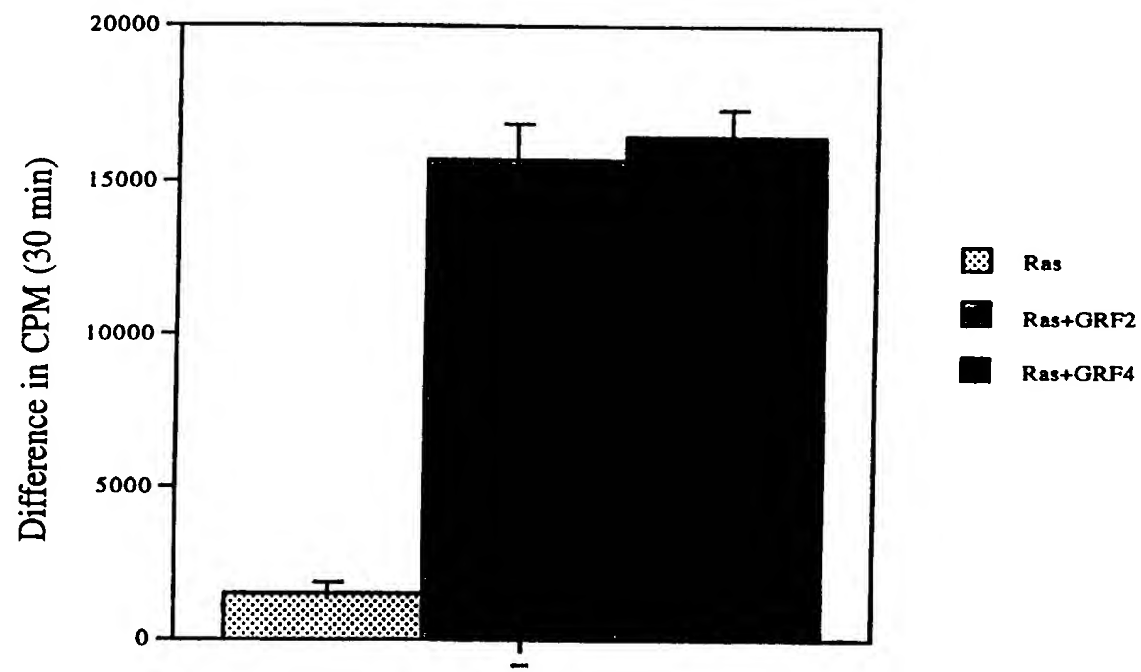


Fig. 15

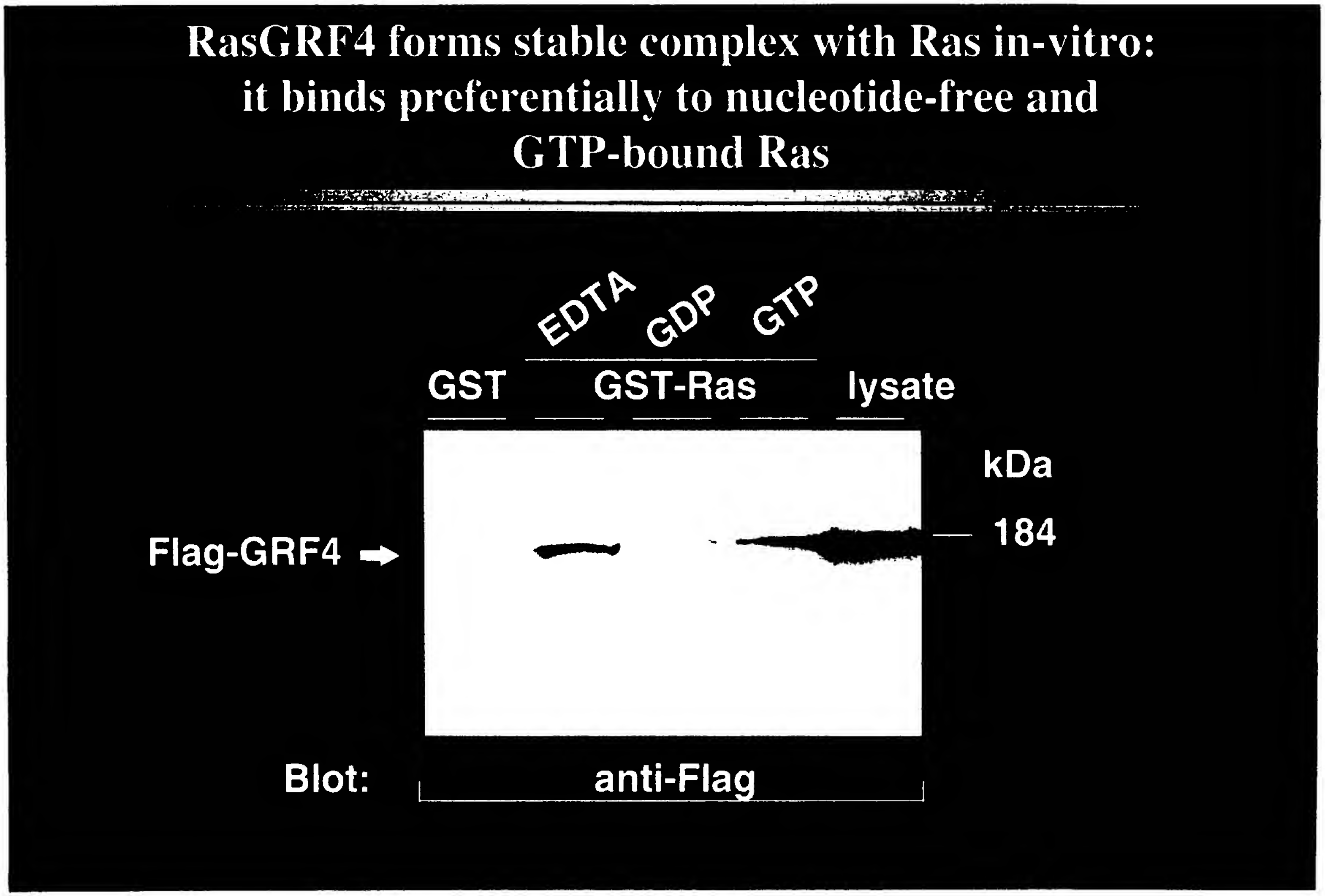

RasGRF4 forms stable complex with Ras in-vitro:
it binds preferentially to nucleotide-free and
GTP-bound Ras
EDTA
GDP
GTP
GST
GST-Ras
lysate
kDa
184
Flag-GRF4
Blot:
anti-Flag

Fig. 16

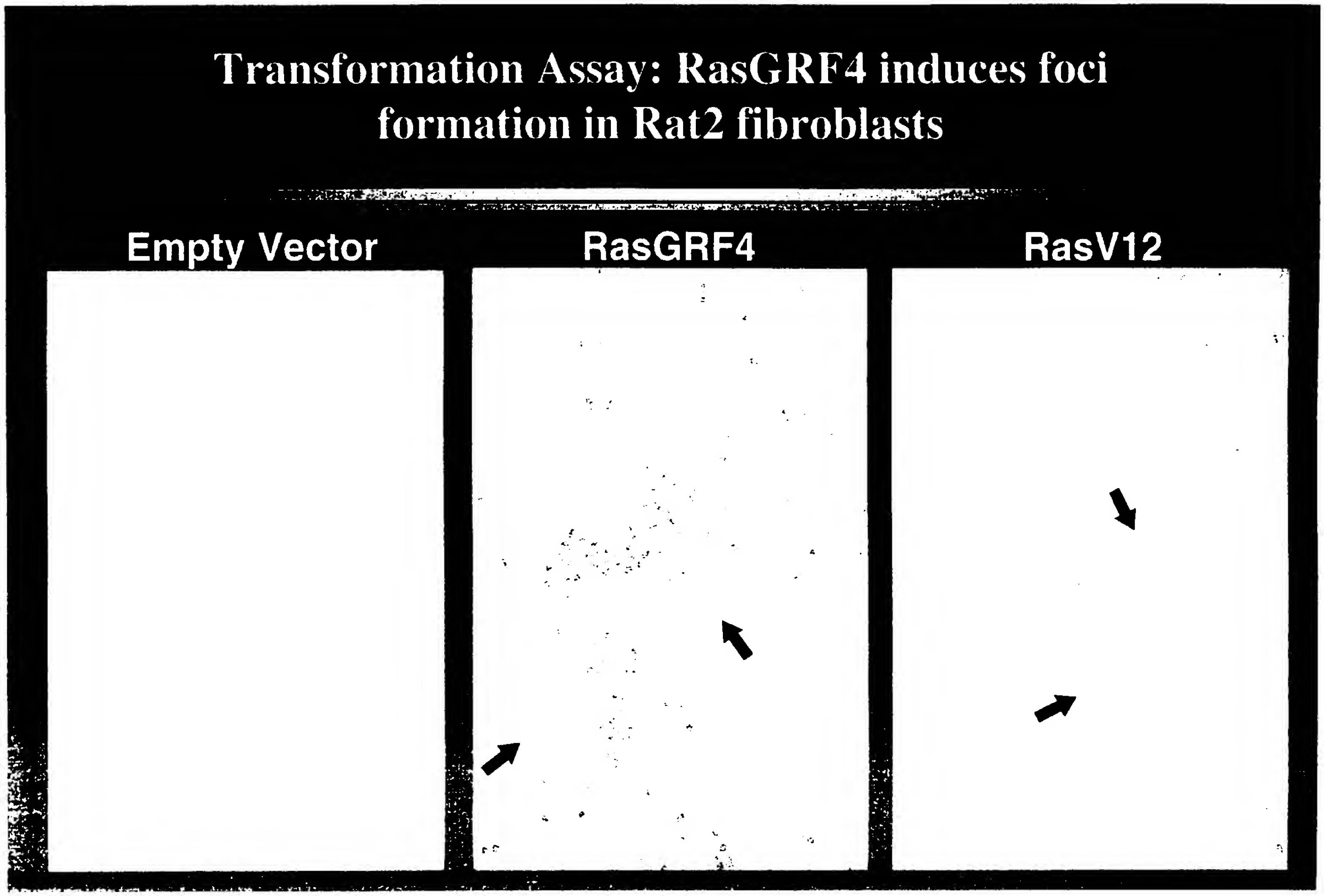

Transformation Assay: RasGRF4 induces foci formation in Rat2 fibroblasts
Empty Vector
RasGRF4
RasV12

Fig. 17

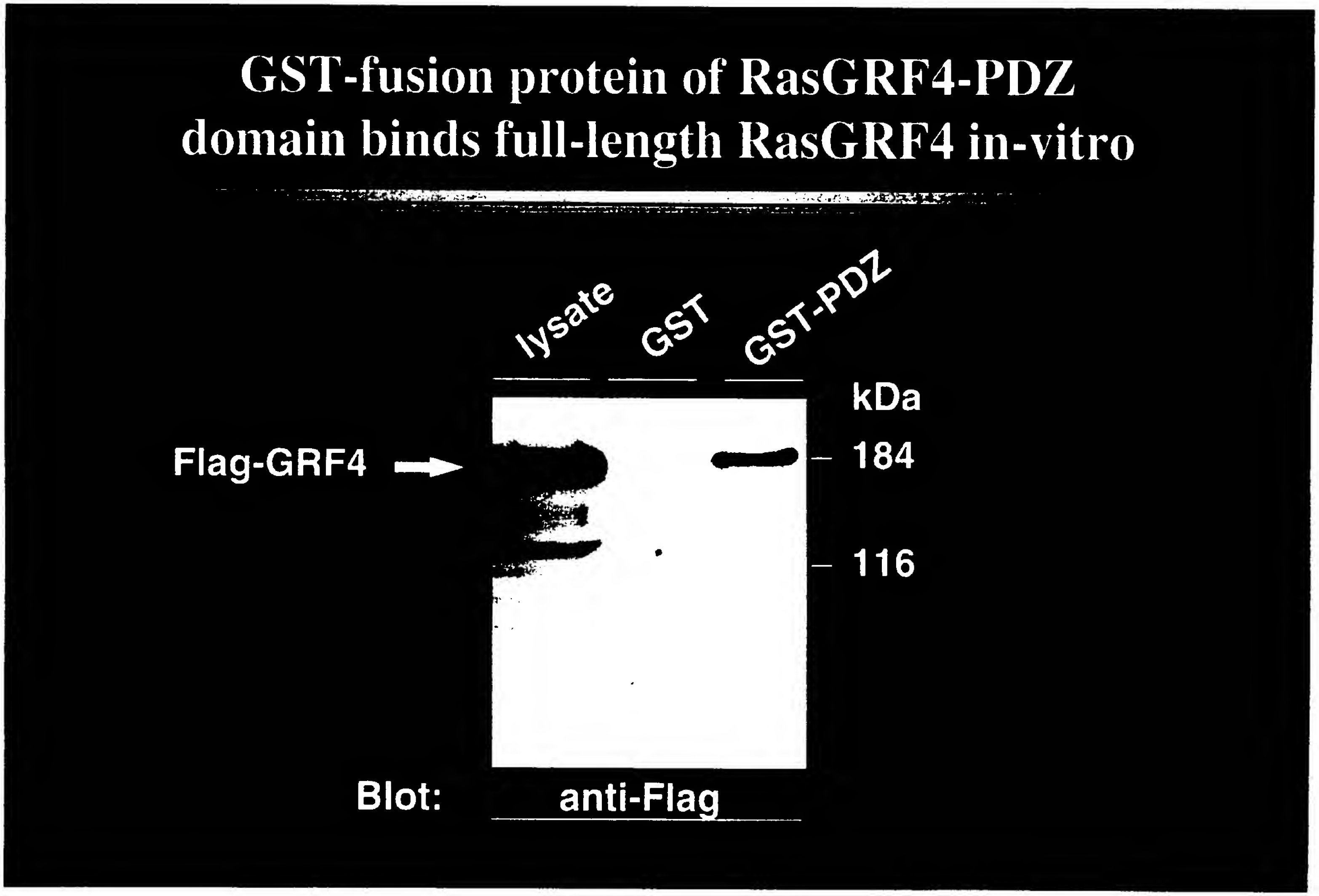
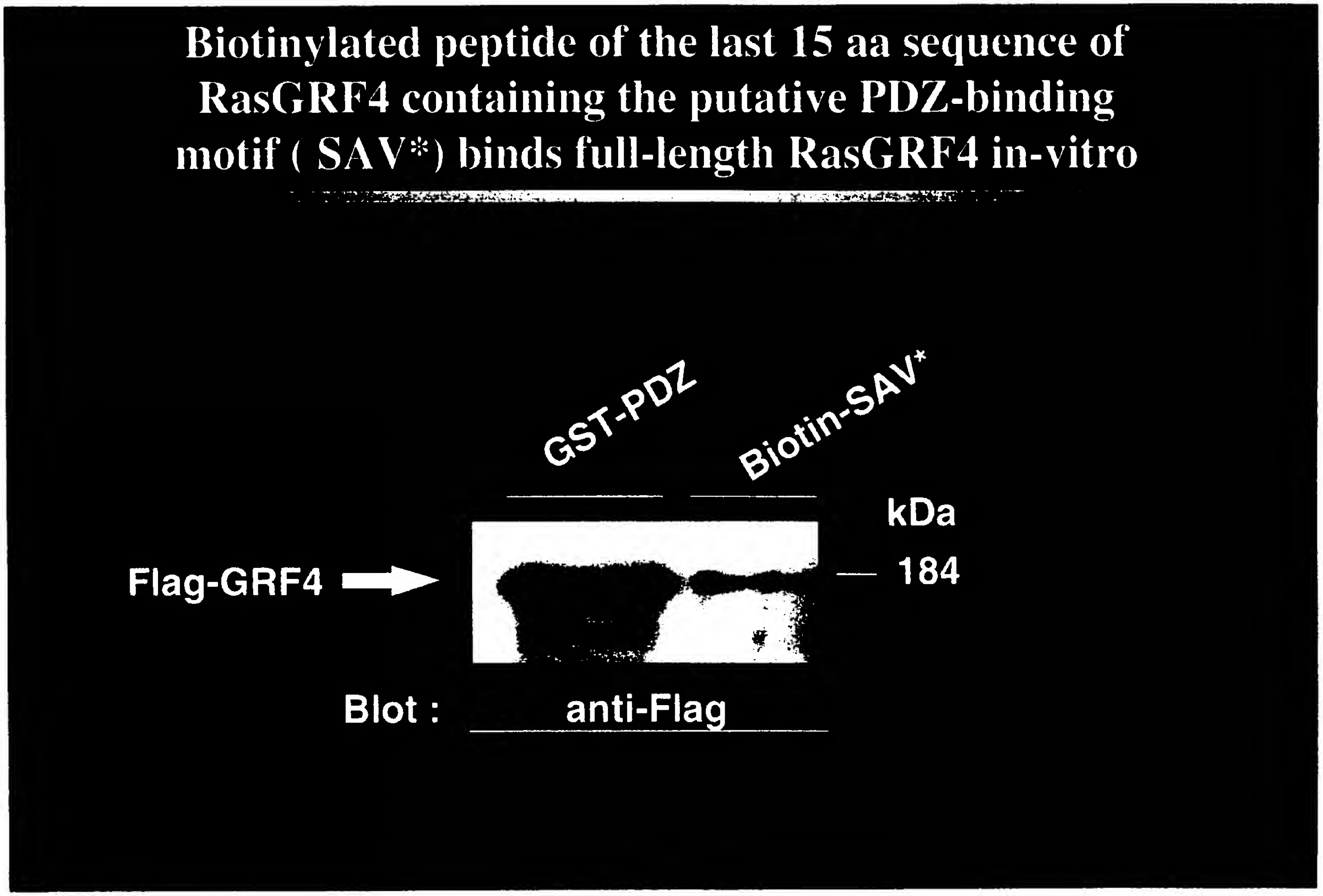

GST-fusion protein of RasGRF4-PDZ
domain binds full-length RasGRF4 in-vitro
lysate
GST
GST-PDZ
kDa
184
116
Flag-GRF4
Blot:
anti-Flag

Fig 18


Biotinylated peptide of the last 15 aa sequence of RasGRF4 containing the putative PDZ-binding motif (SAV*) binds full-length RasGRF4 in-vitro
GST-PDZ
Biotin-SAV*
kDa
184
Flag-GRF4
Blot :
anti-Flag

```
LOCUS       AB002311     6568 bp    mRNA
DEFINITION  Human mRNA for KIAA0313 gene, complete cds.
ACCESSION   AB002311
NID         g2224566
KEYWORDS    KIAA0313.
SOURCE      Homo sapiens male brain cDNA to mRNA, clone_lib:pBluescriptII SK
            plus clone:HG0186.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
```

```
JOURNAL   DNA Res. 4, 141-150 (1997)
FEATURES             Location/Qualifiers
     source          1..6568
                     /organism="Homo sapiens"
                     /clone="HG0186"
                     /clone_lib="pBluescriptII SK plus"
                     /sex="male"
                     /tissue_type="brain"
     CDS             63..4562
                     /gene="KIAA0313"
                     /codon_start=1
                     /db_xref="PID:d1021609"
                     /db_xref="PID:g2224567"
                     /translation="MKPLAIPANHGVMGQQEKHSLPADFTKLHLTDSLHPQVTHVSSS
                     MSGCSITSDSGSSSLSDIYQATESEAGDMDLSGLPETAVDSEDDDDEEDIERASDPLM
                     SRDIVRDCLEKDPIDRTDDDIEQLLEFMHQLPAFANMTMSVRRELCAVMVFAVVERAG
                     TIVLNDGEELDSWSVILNGSVEVTYPDGKAEILCMGNSFGVSPTMDKEYMKGVMRTKV
                     DDCQFVCIAQQDYCRILNQVEKNMQKVEEEGEIVMVKEHRELDRTGTRKGHIVIKGTS
                     ERLTMHLVEEHSVVDPTFIEDFLLTYRTFLSSPMEVGKKLLEWFNDPSLRDKVTRVVL
                     LWVNNHFNDFEGDPAMTRFLEEFENNLEREKMGGHLRLLNIACAAKAKRRLMTLTKPS
                     REAPLPFILLGGSEKGFGIFVDSVDSGSKATEAGLKRGDQILEVNGQNFENIQLSKAM
                     EILRNNTHLSITVKTNLFVFKELLTRLSEEKRNGAPHLPKIGDIKKASRYSIPDLAVD
                     VEQVIGLEKVNKKSKANTVGGRNKLKKILDKTRISILPQKPYNDIGIGQSQDDSIVGL
                     RQTKHIPTALPVSGTLSSSNPDLLQSHHRILDFSATPDLPDQVLKVFKADQQSRYIMI
                     SKDTTAKEVVIQAIREFAVTATPDQYSLCEVSVTPEGVIKQRRLPDQLSKLADRIQLS
                     GRYYLKNNMETETLCSDEDAQELLRESQISLLQLSTVEVATQLSMRNFELFRNIEPTE
                     YIDDLFKLRSKTSCANLKRFEEVINQETFWVASEILRETNQLKRMKIIKHFIKIALHC
                     RECKNFNSMFAIISGLNLAPVARLRTTWEKLPNKYEKLFQDLQDLFDPSRNMAKYRNV
                     LNSQNLQPPIIPLFPVIKKDLTFLHEGNDSKVDGLVNFEKLRMIAKEIRHVGRMASVN
                     MDPALMFRTRKKKWRSLGSLSQGSTNATVLDVAQTGGHKKRVRRSSFLNAKKLYEDAQ
                     MARKVKQYLSNLELEMDEESLQTLSLQCEPATNTLPKNPGDKKPVKSETSPVAPRAGS
                     QQKAQSLPQPQQQPPPAHKINQGLQVPAVSLYPSRKKVPVKDLPPFGINSPQALKKIL
                     SLSEEGSLERHKKQAEDTISNASSQLSSPPTSPQSSPRKGYTLAPSGTVDNFSDSGHS
                     EISSRSSIVSNSSFDSVPVSLHDERRQRHSVSIVETNLGMGRMERRTMIEPDQYSLGS
                     YAPMSEGRGLYATATVISSPSTEELSQDQGDRASLDAADSGRGSWTSCSSGSHDNIQT
                     IQHQRSWETLPFGHTHFDYSGDPAGLWASSSHMDQIMFSDHSTKYNRQNQSRESLEQA
                     QSRASWASSTGYWGEDSEGDTGTIKRRGGKDVSIEAESSSLTSVTTEETKPVPMPAHI
                     AVASSTTKGLIARKEGRYREPPPTPPGYVIGIPITDFPEGHSHPARKPPDYNVALQRSR
                     MVARSSDTAGPSSVQQPHGHPTSSRPVNKPQWHKPNESDPRLAPYQSQGFSTEEDEDE
                     QVSAV"
     gene            63..4562
                     /gene="KIAA0313"
BASE COUNT     1974 a   1400 c   1463 g   1731 t
ORIGIN
        1 cttgccatcg tgagagattg gtacatgatg tgtaaattca gttcagcata tgtttcttca
       61 ttatgaaacc actagcaatc ccagctaacc atggagttat gggccagcag gagaaacact
      121 cacttcctgc agatttcaca aaactgcatc ttactgacag tctccaccca caggtgaccc
      181 acgtttcttc tagccattca ggatgtagta tcactagtga ttctgggagc agcagtcttt
      241 ctgatatcta ccaggccaca gaaagcgagg ctggtgatat ggacctgagt gggttgccag
      301 aaacagcagt ggattccgaa gacgacgacg atgaagaaga cattgagaga gcatcagatc
      361 ctctgatgag cagggacatt gtgagagact gcctagagaa ggacccaatt gaccggacag
```

[SEQ ID NO:2]

Fig 19(a)

```
421 atgatgacat tgaacaactc ttggaattta tgcaccagtt gcctgctttt gccaatatga
481 caatgtcagt gaggcgagaa ctctgtgctg tgatggtgtt cgcagtggtg gaaagagcag
541 ggaccatagt gttaaatgat ggtgaagagc tggactcctg gtcagtgatt ctcaatggat
601 ctgtggaagt gacttatcca gatggaaaag cagaaatact gtgcatggga aatagttttg
661 gtgtctctcc taccatggac aaagaataca tgaaaggagt gatgagaaca aaggtggatg
721 actgccagtt tgtctgcata gcccagcaag attactgccg tattctcaat caagtagaaa
781 agaacatgca aaaagttgaa gaggaaggag agattgttat ggtgaaagaa caccgagaac
841 ttgatcgaac tggaacaaga aagggacaca ttgtcatcaa gggtacctca gaaaggttaa
901 caatgcattt ggtggaagag cattcagtag tagatccaac attcatagaa gactttctgt
961 tgacctatag gacttttctt tctagcccaa tggaagtggg caaaaagtta ttggagtggt
1021 ttaatgaccc gagcctcagg gataaggtta cacgggtagt attattgtgg gtaaataatc
1081 acttcaatga ctttgaagga gatcctgcaa tgactcgatt tttagaagaa tttgaaaaca
1141 atctggaaag agagaaaatg ggtggacacc taaggctgtt gaatatcgcg tgtgctgcta
1201 aagcaaaaag aagattgatg acgttaacaa aaccatcccg agaagctcct ttgcctttta
1261 tcttacttgg aggctctgag aagggatttg gaatctttgt tgacagtgta gattcaggta
1321 gcaaagcaac tgaagcaggc ttgaaacggg gggatcagat attagaagta aatggccaaa
1381 actttgaaaa cattcagctg tcaaaagcta tggaaattct tagaaataac acacatttat
1441 ctatcactgt gaaaaccaat ttatttgtat ttaaagaact tctaacaaga ttgtcagaag
1501 agaaaagaaa tggtgccccc caccttccta aaattggtga cattaaaaag gccagtcgct
1561 actccattcc agatcttgct gtagatgtag aacaggtgat aggacttgaa aaagtgaaca
1621 aaaaaagtaa agccaacact gtgggaggaa ggaacaagct gaaaaagata ctcgacaaga
1681 ctcggatcag tatcttgcca cagaaaccat acaatgatat tgggattggt cagtctcaag
1741 atgacagcat agtaggatta aggcagacaa agcacatccc aactgcattg cctgtcagtg
1801 gaaccttatc atccagtaat cctgatttat tgcagtcaca tcatcgcatt ttagacttca
1861 gtgctactcc tgacttgcca gatcaagtgc taagggtttt taaggctgat cagcaaagcc
1921 gctacatcat gatcagtaag gacactacag caaaggaagt ggtcattcag gctatcaggg
1981 agtttgctgt tactgccacc ccggatcaat attcactatg tgaggtctct gtcacacctg
2041 agggagtaat caaacaaaga agacttccag atcagctttc caaacttgca gacagaatac
2101 aactgagtgg aaggtattat ctgaaaaaca acatggaaac agaaactctt tgttcagatg
2161 aagatgctca ggagttgttg agagagagtc aaatttccct ccttcagctc agcactgtgg
2221 aagttgcaac acagctctct atgcgaaatt ttgaactctt tcgcaacatt gaacctactg
2281 aatatataga tgatttattt aaactcagat caaaaaccag ctgtgccaac ctgaagagat
2341 ttgaagaagt cattaaccag gaaacatttt gggtagcatc tgaaattctc agagaaacaa
2401 accagctgaa gaggatgaag atcattaagc atttcatcaa gatagcactg cactgtaggg
2461 aatgcaagaa ttttaactca atgtttgcaa tcatcagtgg cctaaacctg gcaccagtgg
2521 caagactgcg aacgacctgg gagaaacttc ccaataaata cgaaaaacta tttcaagatc
2581 tccaagacct gtttgatcct tccagaaaca tggcaaaata tcgtaatgtt ctcaatagtc
2641 aaaatctaca acctcccata atccctctat tcccagttat caaaaaggat ctcaccttcc
2701 ttcacgaagg aaatgactca aaagtagacg ggctggtcaa ttttgagaag ctaaggatga
2761 ttgcaaaaga aattcgtcac gttggccgaa tggcttcagt gaacatggac cctgccctca
2821 tgttcaggac tcggaagaag aaatggcgga gtttggggtc tctcagccag ggtagtacaa
2881 atgcaacagt gctagatgtt gctcagacag gtggtcataa aaagcgggta cgtcgtagtt
2941 cctttctcaa tgccaaaaag ctttatgaag atgcccaaat ggctcgaaaa gtgaagcagt
3001 acctttccaa tttggagcta gaaatggacg aggagagtct tcagacatta tctctgcagt
3061 gtgagccagc aaccaacaca ttgcctaaga atcctggtga caaaaagcct gtcaaatccg
3121 agacctctcc agtagctcca agggcagggt cacaacagaa agctcagtcc ctgccacagc
3181 cccagcagca gccaccacca gcacataaaa tcaaccaggg actacaggtt cccgccgtgt
3241 ccctttatcc ttcacggaag aaagtgcccg taaaggatct cccacctttt ggcataaact
3301 ctccacaagc tttaaaaaaa attctttctt tgtctgaaga aggaagtttg gaacgtcaca
3361 agaaacaggc tgaagataca atatcaaatg catcttcgca gctttcttct cctcctactt
3421 ctccacagag ttctccaagg aaaggctata ctttggctcc cagtggtact gtggataatt
3481 tttcagattc tggtcacagt gaaatttctt cacgatccag tattgttagc aattcgtctt
3541 ttgactcagt gccagtctca ctgcacgatg agaggcgcca gaggcattct gtcagcatcg
3601 tggaaacaaa cctagggatg ggcaggatgg agaggcggac catgattgaa cctgatcagt
3661 atagcttggg gtcctatgca ccaatgtccg agggccgagg cttatatgct acagctacag
3721 taatttcttc tccaagcaca gaggaacttt cccaggatca gggggatcgc gcgtcacttg
3781 atgctgctga cagtggccgt gggagctgga cgtcatgctc aagtggctcc catgataata
3841 tacagacgat ccagcaccag agaagctggg agactcttcc attcgggcat actcactttg
3901 attattcagg ggatcctgca ggtttatggg catcaagcag ccatatggac caaattatgt
3961 tttctgatca tagcacaaag tataacaggc aaaatcaaag tagagagagc cttgaacaag
4021 cccagtcccg agcaagctgg gcgtcttcca caggttactg gggagaagac tcagaaggtg
4081 acacaggcac aataaagcgg aggggtggaa aggatgtttc cattgaagcc gaaagcagta
4141 gcctaacgtc tgtgactacg gaagaaacca agcctgtccc catgcctgcc cacatagctg
4201 tggcatcaag tactacaaag gggctcattg cacgaaagga gggcaggtat cgagagcccc
4261 cgcccacccc tcccggctac attggaattc ccattactga ctttccagaa gggcactccc
4321 atccagccag gaaaccgccg gactacaacg tggcccttca gagatcgcgg atggtcgcac
4381 gatcctccga cacagctggg ccttcatccg tacagcagcc acatgggcat cccaccagca
4441 gcaggcctgt gaacaaacct cagtggcata aaccgaacga gtctgacccg cgcctcgccc
4501 cttatcagtc ccaagggttt tccaccgagg aggatgaaga tgaacaagtt tctgctgttt
4561 gaggcacaga cttttctgga agcagagcga gccacctgaa aggagagcac aagaagacgt
4621 cctgagcatt ggagccttgg aactcacatt ctgaggacgg tggaccagtt tgcctccttc
4681 cctgccttaa aagcagcatg ggcttcttc tcccccttctt cctttcccct ttgcatgtga
4741 aatactgtga agaaattgcc ctggcacttt tcagactttg ttgcttgaaa tgcacagtgc
4801 agcaatcttc gagctcccac tgttgctgcc tgccacatca cacagtatca ttccaaattc
4861 caagatcatc acaacaagat gattcactct ggctgcactt ctcaatgcct ggaaggattt
4921 tttttaatct tccttttaga tttcaatcca gtcctagcac ttgatctcat tgggataatg
4981 agaaaagcta gccattgaac tacttggggc ctttaaccca ccaaggaaga caaagaaaaa
5041 caatgaaatc ctttgagtac agtgcttgtc cacttgttta caatgtcctc cttttaaaaa
5101 aaaaaatgag tttaaagatt ttgttcagag agtaaatata tatccattta atgattacag
```

5161 tattatttta aaccttaagt agggttgcca gcctggtttc tgaaaaacca aatatgccgg
5221 acagggtgtg gccacaccaa gaagacggga agacctggct tgtgaccctg gcttcccatg
5281 tccttctggt ctcacccgcg aagtgcccta tcctggaagt atgaaatgtt agccaattaa
5341 taccaagaca cctcatctgc tccttcccca gtggatgggg ttcttctgta aaactgtttg
5401 cacatggcca ggggagggaa ctaggaccct tgtgtcctgt ctgagcctta tggaggcagg
5461 acggtgtcat tggcggatgt gtcctgctcc attgagatgg atggcaaacc ccatttttaa
5521 gttatatttc tttgattttt gttaatttag aggtgtaggt tttgtttttt gtttttttgtt
5581 ttttttttaag agaaacattt ataactggat agcattgcag tgaaagcagc ttgggatgtt
5641 ggagctaatg ccagctgttt atactgctct ttcaagacag cctcccttta ttgaattggc
5701 attagggaat aaacaagcct ttaaacgtga taaaagatca aaaacctggt tagacatgcc
5761 agcctttgca aggcaggtta gtcaccaaag actaacctcc aagtggcttt atggacgctg
5821 catatagaga aggcctaagt gtagcaacca tctgctcaca gctgctatta accctataat
5881 gactgaaatg acccctccac tctatttttg tgttgttttg cacagactcc ggaaaagtga
5941 aggctgccaa tctgagtagt actcaaatgt gaggaactgc tggtcttgga ttttttttcc
6001 attaaattca gctgatcata ttgatcagta gataaacgta aatagcttca aattttaaaa
6061 gtggaattgc agtgtttttt cactgtatca aacaatgtca gtgctttatt taataattct
6121 cttctgtatc atggcatttg tctacttgct tattacattg tcaattatgc atttgtaatt
6181 ttacatgtaa tatgcattat ttgccagttt tattatatag gctatggacc tcatgtgcat
6241 atagaaagac agaaatctag ctctaccaca agttgcacaa atgttatcta agcattaagt
6301 aattgtagaa cataggactg ctaatctcag ttcgctctgt gatgtcaagt gcagaatgta
6361 caattaactg gtgatttcct catacttttg atactacttg tacctgtatg tcttttagaa
6421 agacattggt ggagtctgta tcccttttgt atttttaata caataattgt acatattggt
6481 tatatttttg ttgaagatgg tagaaatgta ctatgtttat gcttctacat ccagtttgta
6541 caagctggaa aataaataaa tataacat

[SEQ ID NO: 1]

DNA Strider™ 1.2

Fig 19(b)

7.7 2.8 acgt strider 800 bp -> 3-phase Translation

DNA sequence 801 b.p. ACTAAAGGGAAC ... TGCCACGTCACG linear

```
1/1                                          31/11
ACT AAA GGG AAC AAA AGC TGG AGC TCC ACC GCG GTG GCG GCC GCT CTA GAA CTA GTG GAT
T   K   G   N   K   S   W   S   S   T   A   V   A   A   A   L   E   L   V   D
 L   K   G   T   K   A   G   A   P   P   R   W   R   P   L   *   N   *   W   I
  *   R   E   Q   K   L   E   L   H   R   G   G   G   R   S   R   T   S   G   S
61/21                                        91/31
CCC CCG GGC TGC AGG AAT TCA AGC GGT GGG AAG GAT GTC TCC GCT GAG GCA GAG AGC AGC
P   P   G   C   R   N   S   S   G   G   K   D   V   S   A   E   A   E   S   S
 P   R   A   A   G   I   Q   A   V   G   R   M   S   P   L   R   Q   R   A   A
  P   G   L   Q   E   F   K   R   W   E   G   C   L   R   *   G   R   E   Q   Q
121/41                                       151/51
AGC ATG GTG CCC GTG ACT ACA GAG GAA GCC AAA CCT GTC CCT ATG CCT GCC CAC ATA GCT
S   M   V   P   V   T   T   E   E   A   K   P   V   P   M   P   A   H   I   A
 A   W   C   P   *   L   Q   R   K   P   N   L   S   L   C   L   P   T   *   L
  H   G   A   R   D   Y   R   G   S   Q   T   C   P   Y   A   C   P   H   S   C
181/61                                       211/71
GTG ACG CCG AGC ACT ACC AAG GGA CTC ATC GCA CGG AAG GAA GGC AGG TAC CGG GAG CCG
V   T   P   S   T   T   K   G   L   I   A   R   K   E   G   R   Y   R   E   P
 *   R   R   A   L   P   R   D   S   S   H   G   R   K   A   G   T   G   S   R
  D   A   E   H   Y   Q   G   T   H   R   T   E   G   R   Q   V   P   G   A   A
241/81                                       271/91
CCT CCC ACA CCT CCA GGC TAC GTG GGC ATC CCC ATT GCC GAT TTC CCA GAA GGG CCT TGC
P   P   T   P   P   G   Y   V   G   I   P   I   A   D   F   P   E   G   P   C
 L   P   H   L   Q   A   T   W   A   S   P   L   P   I   S   Q   K   G   L   A
  S   H   T   S   R   L   R   G   H   P   H   C   R   F   P   R   R   A   L   P
301/101                                      331/111
CAC CCG GCC AGG AAG CCC CCG GAT TAC AAC GTG GCC CTG CAG CGG TCC CGC ATG GTG GCA
H   P   A   R   K   P   P   D   Y   N   V   A   L   Q   R   S   R   M   V   A
 T   R   P   G   S   P   R   I   T   T   W   P   C   S   G   P   A   W   W   H
  P   G   Q   E   A   P   G   L   Q   R   G   P   A   A   V   P   H   G   G   T
361/121                                      391/131
CGG CCC ACT GAG GCC CCG GCA CCG GGC CAG ACG CCG CCT GCA GCC GCA GCC AGC CGG CCG
R   P   T   E   A   P   A   P   G   Q   T   P   P   A   A   A   A   S   R   P
 G   P   L   R   P   R   H   R   A   R   R   R   L   Q   P   Q   P   A   G   R
  A   H   *   G   P   G   T   G   P   D   A   A   C   S   R   S   Q   P   A   G
421/141                                      451/151
GGC AGC AAG CCA CAG TGG CAC AAG CCC AGC GAC GCA GAC CCA CGC CTC GCG CCC TTC CAG
G   S   K   P   Q   W   H   K   P   S   D   A   D   P   R   L   A   P   F   Q
 A   A   S   H   S   G   T   S   P   A   T   Q   T   H   A   S   R   P   S   S
  Q   Q   A   T   V   A   Q   A   Q   R   R   R   P   T   P   R   A   L   P   A
481/161                                      511/171
CCG CAG GCT TCG CAG GAG CGG AGG AGG ACG AAG ATG AAC AAG TGT CTG CTG TTT GAG GCG
P   Q   A   S   Q   E   R   R   R   T   K   M   N   K   C   L   L   F   E   A
 R   R   L   R   R   S   G   G   G   R   R   *   T   S   V   C   C   L   R   R
  A   G   F   A   G   A   E   E   D   E   D   E   Q   V   S   A   V   *   G   A
541/181                                      571/191
CAG GCT CCT TGA TCC ACA GTG AGC CAC CCA AAG GAG AGC ACA AGA AGA CGT CCC AAG CCT
Q   A   P   *   S   T   V   S   H   P   K   E   S   T   R   R   R   P   K   P
 R   L   L   D   P   Q   *   A   T   Q   R   R   A   Q   E   D   V   P   S   L
  G   S   L   I   H   S   E   P   P   K   G   E   H   K   K   T   S   Q   A   L
601/201                                      631/211
TGG AGC CTT GGC ACG CAC ATC TGA GGA TGG TGG ACC AGT TTG CCT CCT TCC CTG CCT TAA
W   S   L   G   T   H   I   *   G   W   W   T   S   L   P   P   S   L   P   *
 G   A   L   A   R   T   S   E   D   G   G   P   V   C   L   L   P   C   L   K
  E   P   W   H   A   H   L   R   M   V   D   Q   F   A   S   F   P   A   L   K
661/221                                      691/231
AGC AGC ATG GGG CTT CTT CTC CCC TTC TTC CTT TCC CCT TTG CAT GTG AAA TAC TGT GAA
S   S   M   G   L   L   L   P   F   F   L   S   P   L   H   V   K   Y   C   E
 A-  A   W   G   F   F   S   P   S   S   F   P   L   C   M   *   N   T   V   K
  Q   H   G   A   S   S   P   L   L   P   P   F   F   A   C   E   I   L   *   R
```

7.7 2.8 acgt strider 800 bp -> 3-phase Translation

```
721/241                                  751/251
GAA ATT GCC CTG GCA CTT TGC AGA CTT GTT GCT TGA AAT GCA CAG CCC AGC AGC CCC TGA
 E   I   A   L   A   L   C   R   L   V   A   *   N   A   Q   P   S   S   P   *
  K   L   P   W   H   F   A   D   L   L   L   E   M   H   S   P   A   A   P   E
   N   C   P   G   T   L   Q   T   C   C   L   K   C   T   A   Q   Q   P   L   S
781/261
GCT GCT GCC TGC CAC GTC ACG [SEQ ID NO:3]
 A   A   A   C   H   V   T  [SEQ ID NO 4]
  L   L   P   A   T   S     [SEQ ID NO:5]
   C   C   L   P   R   H    [SEQ ID NO:6]
```

The PDZ domain (but not the C terminal SxV motif) is required for plasma membrane localization of ras GRF4

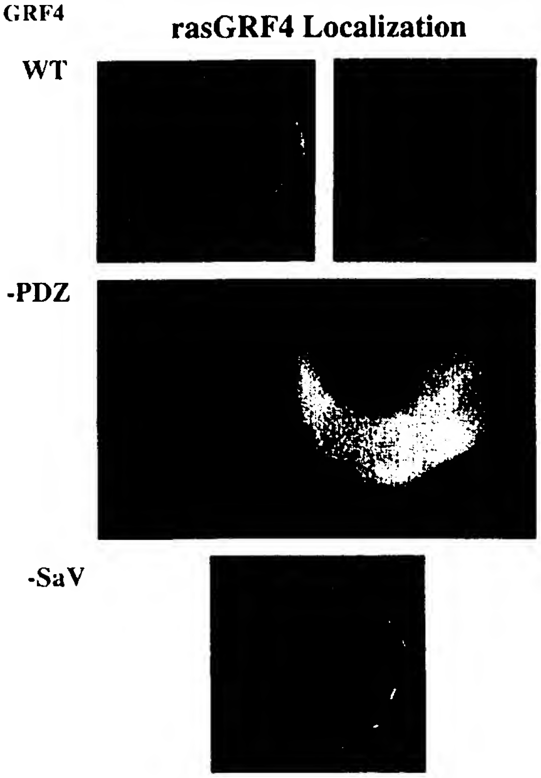


Fig. 20